PERFORMANCE OF THE

DEFENSE ACQUISITION SYSTEM

2014 ANNUAL REPORT

JUNE 13, 2014

Approved for public release; distribution unlimited

			Form Approved OMB No. 0704-0188

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **13 JUN 2014**	2. REPORT TYPE	3. DATES COVERED **00-00-2014 to 00-00-2014**

4. TITLE AND SUBTITLE **Performance of the Defense Acquisition System. 2014 Annual Report**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Under Secretary of Defense, Acquisition, Technology, and Logistics (USD[AT&L],3000 Defense Pentagon,Washington,DC,20301-3000**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**126**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

FOREWORD

Inadequate incentives mean dissolution, or changes of organization purpose, or failure to cooperate. Hence, in all sorts of organizations the affording of adequate incentives becomes the most definitely emphasized task in their existence. It is probably in this aspect of executive work that failure is most pronounced.

—Chester Barnard[1]

By human nature, performance is incentive-driven. Organizations and individuals combine often competing incentives as they take actions and set priorities.

In our second annual report on acquisition performance, we focus on incentives—particularly those from contract types and profits or fees—in addition to updating our prior analysis for which recent data might affect the statistical results. Most of the development and production on acquisition programs is conducted by industry under contract to the government. Therefore, we examine various incentive techniques to see how effective they are at driving cost, schedule, and technical performance.

As in our first report, we emphasize objective data analysis as we examine the effects of statutes, policies, and tradecraft on the outcomes of acquisition. Often our data are limited in the number of cases that report or because significant manual effort is required to structure the data for analysis. For these reasons, we continue to rely on statistical techniques to identify trends and gauge the consistency (and thus broader applicability) of our findings.

As our body of analysis grows, we are finding more practical insights from their results. Of particular note, this year's report shows that the prevalent debate of cost-type versus fixed-price contracting is misguided. The real issue is how effective the incentives are for each contract type within each of those groupings. There are cost-type contracts that are very effective at controlling cost while others are not. Fixed-price contracts are effective at controlling cost—but some types do not share those savings with the government, yielding higher margins (sometimes spectacularly so) and higher prices for our taxpayers.

Moreover, the label "fixed price" can be misunderstood: Prices on fixed-price contracts are only "fixed" if the contractual work content and deliverables remain fixed; such contracts can be (and often are) easily modified to handle unexpected technology gaps, engineering issues or shifting threats, leading to cost growth. At times fixed price vehicles can be virtually

---

[1]Barnard (1938, p. 139), as quoted in Laffont and Martimort (2001).

indistinguishable from cost plus vehicles, as was the case with the Air Force's canceled Expeditionary Combat Support System (ECSS).
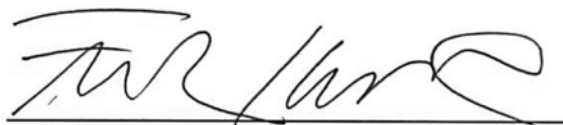
These findings do not, however, dictate "one size fits all" policies. Defense acquisition is complicated and varying. There are no simple "schoolhouse" solutions that should be mandated absent the particulars of each acquisition in question. These findings do, however, inform our individual program decisions and provide fresh insights into what generally works in what circumstances and why. This is critical to improving our performance: We must empower, encourage, and train our workforce to think—not dictate a cookbook that workforce members blindly follow.

This report continues to reflect results from ongoing Department of Defense (DoD) compliance with the Improve Acquisition Act of 2010 and the earlier Weapon Systems Acquisition Reform Act on performance assessments of the defense acquisition system. While similarly motivated, our efforts go beyond the specifics of these acts to seek additional insights for improving the performance of the defense acquisition system.

In addition, this study fulfills ongoing requests from the Office of Management and Budget for an evidence-based analytic study on acquisition performance.

The release of this year's annual performance report has been timed so that we can include analysis of the latest Selected Acquisition Reports (released 45 days after the President's Budget Request to Congress).

In these times of extreme budget pressures and uncertainty, combined with evolving and increasing national security threats, particularly threats to our technological superiority, improving the performance of the defense acquisition system is essential for the DoD. We must ensure that our acquisition professionals have the knowledge they need to incentivize industry and to control cost, schedule, and performance. This report is one of many steps we are taking to achieve that goal.

The Honorable Frank Kendall
Under Secretary of Defense for
Acquisition, Technology, and Logistics

ORGANIZATION OF THE REPORT

Chapter 1 (this chapter) provides background material on acquisition, spending levels and trends, and general perspectives on measuring institutional performance to set the stage for the analysis presented in subsequent chapters.

Chapter 2 serves as an overview on how well we are doing. It analyzes the performance outcomes of our acquisition institutions from a variety of perspectives, including performance: DoD-wide, by commodity type, from contract- and program-level views, by military department, and by prime contractor. This chapter builds on the results from last year's USD(AT&L) report, updating some analysis and providing new results using different datasets. To a large extent, this chapter only measures our performance and does not explain broader policy implications, which are discussed in the next chapter.

Chapter 3 goes beyond the performance analysis in Chapter 2 and discusses actionable insights and lessons learned from analyzing specific practices. This year we focus on the effectiveness of incentives, including contract type and margins.

Chapter 4 provides closing observations, continuing analytic issues and challenges, and objectives for future analysis and reports.

Appendix A provides details on the statistical methods employed in the reported analyses.

(This page was left blank intentionally)

Contents

ACKNOWLEDGEMENTS

(This page was left blank intentionally)

1. THE ACQUISITION LANDSCAPE

Our acquisition system—its institutions, offices, laboratories, workforce, managers, executives, and industrial partners—conducts research, provides a wide range of services, develops and produces new goods, and sustains their capabilities for warfighters and other operators. The performance of that system is measured relative to its outputs and outcomes of interest. Identifying internal policies, processes, workforce, and management capabilities that bear positively or negatively on those measures requires data and analysis to avoid speculative or cyclical policy choices based on current conventional wisdom or untested hypotheses.

FOLLOWING THE MONEY: SPEND ANALYSIS

The defense acquisition system acquires goods and services to support our military forces both now and in the future, while fulfilling our responsibility to be efficient and to avoid waste, fraud, and abuse of taxpayer dollars. The department budgets and accounts for expenditures in various ways—each of which provides useful perspective on what the largest expenditures are for.

Spending by Comptroller Budget Accounts

Broken down by the Comptroller's budget accounts, the President's FY2015 base budget (PB15) requests $90.4 billion for Procurement and $63.5 billion for Research, Development, Test and Evaluation (RDT&E)—see Figure 1-1. Of this $153.9 billion, 45 percent ($69.6 billion) is for programs that have been designated as Major Defense Acquisition Programs (MDAPs), which provide the bulk of the readily available program data for analysis in this year's report. In addition, the PB15 also requests $198.7 billion for Operations and Maintenance (O&M) and $135.2 billion for Military Personnel (MilPer). A sizable portion of O&M is also spent on contracts for goods and services; thus, this portion is also part of the defense acquisition system. Supplemental funding for Overseas Contingency Operations (OCO) is not included in PB15 figures.

Figure 1-2 shows how defense budget accounts have changed over time and compares these to PB15. Analysis of DoD budgets since 1980 found that the total budget oscillates in a sinusoidal pattern that repeats about every 24 years (plus inflationary changes and noise). The current budget is on the second half of the falling portion of the general pattern. Future budgets, of course, are hard to predict, but these patterns show some structure to recent budgetary ups and downs.

Figure 1-1. Defense Budget Breakouts in the 2015 President's Budget Request

by Budget Account

by Military Department
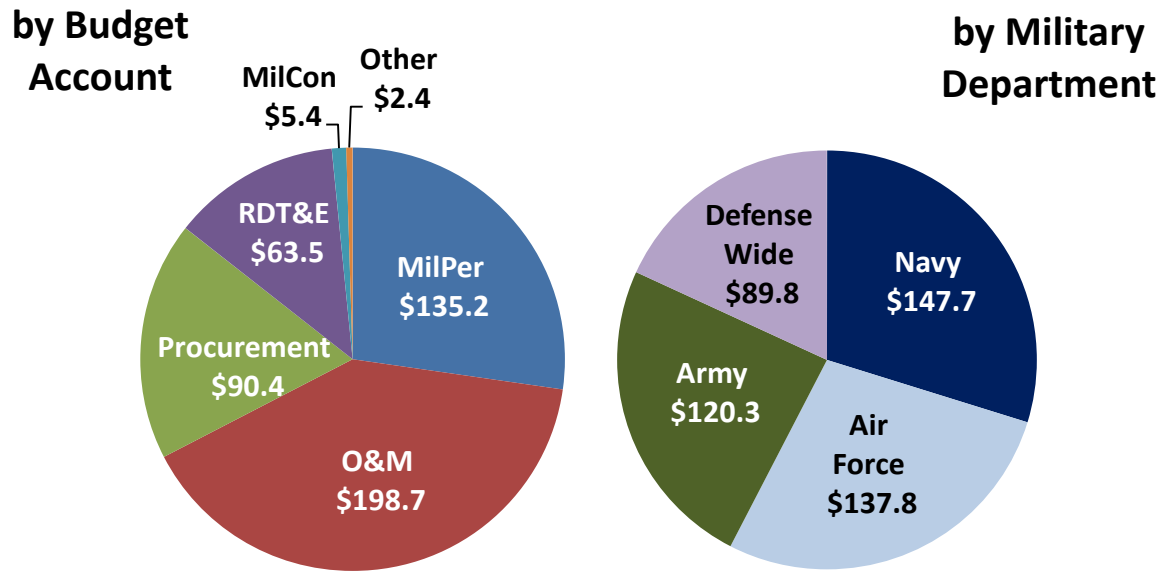
NOTE: Budget amounts are in billions of then-year (unadjusted) dollars. OCO is not included. "MilCon" is Military Construction.

Figure 1-2. Defense Budget Accounts: Historical and PB15 (FY1962–FY2019)
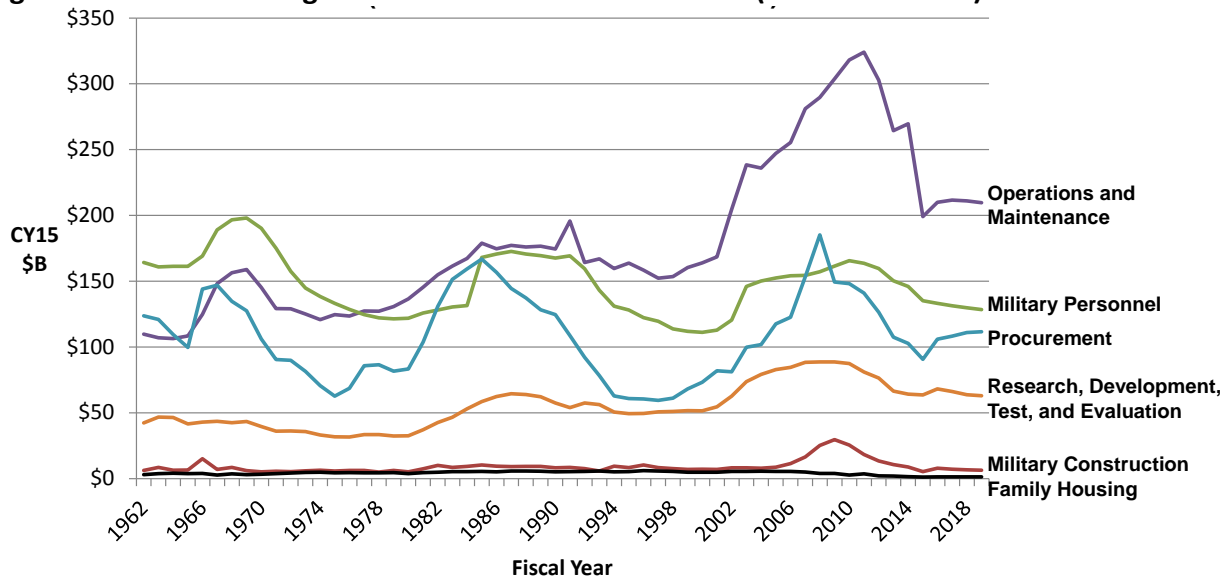
NOTE: OCO is included in fiscal year budgets before FY2014 but not in the current fiscal year (2014) or in the FY2015 President's Budget figures (FY 2015–2019). Budget amounts are adjusted for inflation and reported in billions of calendar year 2015 dollars (CY15$B).

Operation and Maintenance Budgets

Since the mid-1970s, O&M has been the largest part of the DoD's budget (see Figure 1-2). There is no indication this will change. A line connecting the low points in O&M before and after each surge shows this upward trend. As is typical during wartime, O&M increased dramatically to support operations in Afghanistan and Iraq, but that only explains part of the trend in O&M.

In addition to wartime surges, rising O&M may reflect cost increases in a number of things such as medical care, operation and repair of military facilities, and civilian pay differentials. Operation and support (O&S) of weapon systems may also be increasing for various reasons, including increased maintenance costs for more sophisticated systems and increased contractor support. However, O&S may be decreasing in certain situations (e.g., when replacing outdated systems that are hard to maintain). To better understand these factors and their relative contributions to rising O&M costs, the Department has been working to increase the fidelity of our O&S estimates and measures (e.g., by improving how we measure O&S and by improved tracking of O&S by component and vehicle). Preliminary insights on O&M acquisition costs are discussed in this report and are the subject of continued analysis.

RDT&E Budgets

Figure 1-3 shows the breakdown of RDT&E budgets by budget activity. Complete data are readily available only since 2000, but this provides a useful picture of how these accounts have fared during the recent budgetary surge and subsequent decline. Here the science and technology (S&T) accounts (6.1–6.3) are relatively flat or returned to their pre-2001 levels. Accounts for Advanced Component Development and Prototypes (6.4) and Operational Systems Development (6.7, for existing systems) are projected to come down from their peak but stay above the pre-2001 level. System Development and Demonstration (6.5) for new systems in the DoD's product "pipeline" is projected to decline well below 2001 levels. Under sequestration further cuts to R&D would occur.

Figure 1-3. Recent and Projected RDT&E Budgets as of PB2015 (FY2000–FY2019)

NOTE: OCO is included in fiscal year budgets before FY2014 but not in the current fiscal year (2014) or in the FY2015 President's Budget figures (FY 2015–2019). Budget amounts are adjusted for inflation and reported in billions of calendar year 2015 dollars (CY15$B).
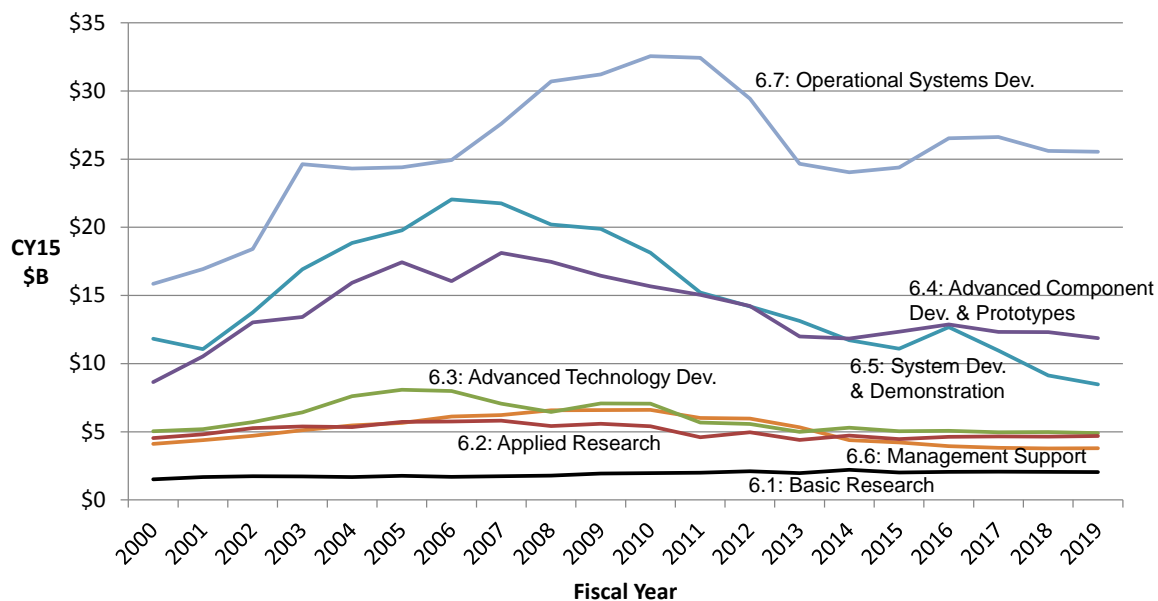
Contractual Spending by Product Service Code Portfolios

Much of what we acquire comes through contracts with industry. Thus, a different way to understand what the defense acquisition system acquires is to examine contract obligations by type as opposed to budget account.

The contracting community uses a categorization called *product service codes* (PSCs) to track what is procured under federal contracts.[2] The Federal Procurement Data System—Next Generation (FPDS-NG) records PSCs for every contract obligation worth at least $3,000, so this taxonomy affords us a way to quickly look across all DoD external (contracted) spending.

At the top level, spending (obligations in this case) is split between *products* (also referred to as *supplies and equipment* [S&E]) and *contracted services*.[3] Figure 1-4 shows that in FY13, just over half (52 percent) of contract obligations were for contracted services. Some caution is warranted, however. While the acquisition community generally considers RDT&E as part of developing a physical system, contract PSCs identify research and development (R&D) as a service (i.e., it is dominated by tasks that do not produce physical end items of supply). Also contract obligations often include multiple types of work, but only one PSC is reported per obligation.

Figure 1-5 shows a further breakdown of all DoD contract obligations by portfolio groupings we developed of these PSCs. Here we see that some contracting has remained relatively flat over the years while others are declining with the recent budget cutbacks.

Figure 1-4. Total DoD Contract Obligations Split Between Goods and Services (FY2013)

FY13
(BY15$, billion)

Supplies & equipment $152.6 48%

Contracted services $166.1 52%

---
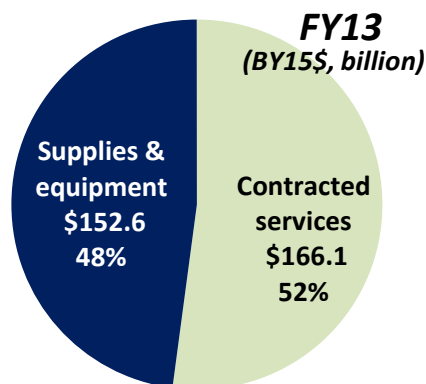
[2]See the Product and Service Codes Manual published by the U.S. General Services Administration (2011). PSCs are recorded in the FPDS-NG to categorize what each federal contract acquires.

[3]The Federal Acquisition Regulation defines a *service contract* as "a contract that directly engages the time and effort of a contractor whose primary purpose is to perform an identifiable task rather than to furnish an end item of supply" (see FAR, sec. 37.101). Because the DoD often refers to the military departments (i.e., Army, Navy, and Air Force) as "Services," this report capitalizes "Services" when referring to military departments but uses lower-case "services" when referring to contracted services.

Figure 1-5. Total DoD Contract Obligations by Portfolio Group (FY2008–2013)
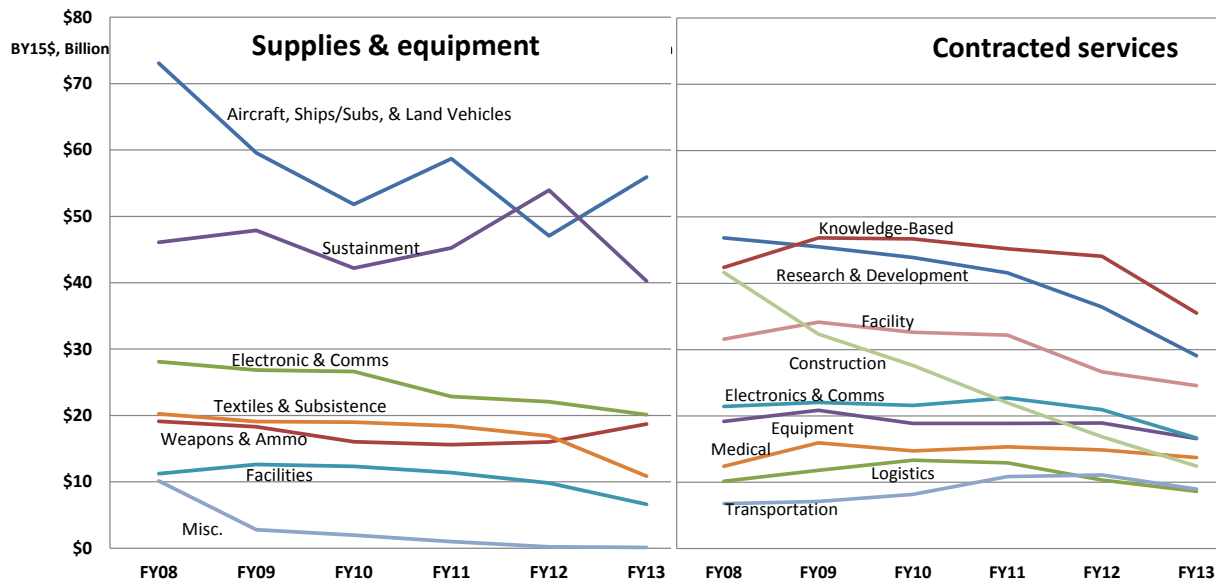
NOTE: FPDS-NG data on all DoD obligations (including OCO). Obligations for contracted services in this period ranged from 57 percent in FY10 to 52 percent in FY13. All numbers are in billions of adjusted BY15 dollars.

PHASES OF ACQUISITION ASSESSED

This report assesses how our institutions perform primarily using existing oversight data aggregated to look for broader performance and trends. Because of their size and the risks involved, most readily available data are on MDAPs and their measurable outcomes. These data include both goods (i.e., production of the weapon systems themselves) and services (i.e., development and testing of those weapon systems).

Figure 1-6 depicts the entire defense acquisition system life cycle[4] and the portion where we currently have the best data for analysis—namely, for development and production up to full operational capability (FOC). While we have some data that reflect partially on the performance in other phases (e.g., early research, analysis of alternatives [AoAs], early risk reduction, and sustainment), operation and support are reflected at best by early estimates. These other phases will be expanded in subsequent versions of this report as we improve data access, quality, and availability.

---

[4] See https://dag.dau.mil/Pages/acqframework.aspx for a description of the life-cycle phases of the Defense Acquisition Management System framework.

Figure 1-6. The Defense Acquisition System Life Cycle
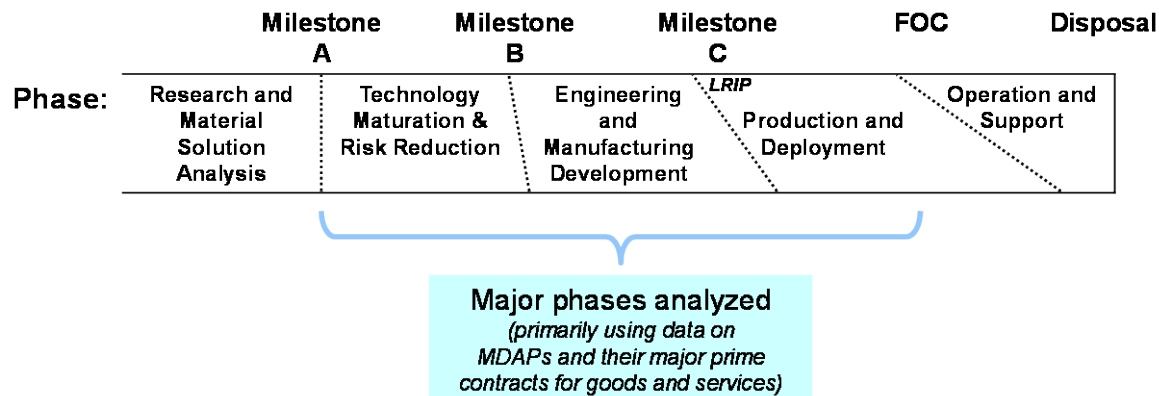
FOC = Full Operational Capability; LRIP = low-rate initial production

ON MEASURING PERFORMANCE

Institutional performance is all about getting value. Value to the Department stems from the immediate benefits (i.e., technical performance) of the goods and services acquired in a responsive time (schedule) compared to the costs to the taxpayer. Hence, measures of cost, schedule, and performance serve as the basis for measuring the effectiveness of the acquisition system in converting inputs to outputs (see Figure 1-7).

Figure 1-7. Output Measures for the Defense Acquisition System
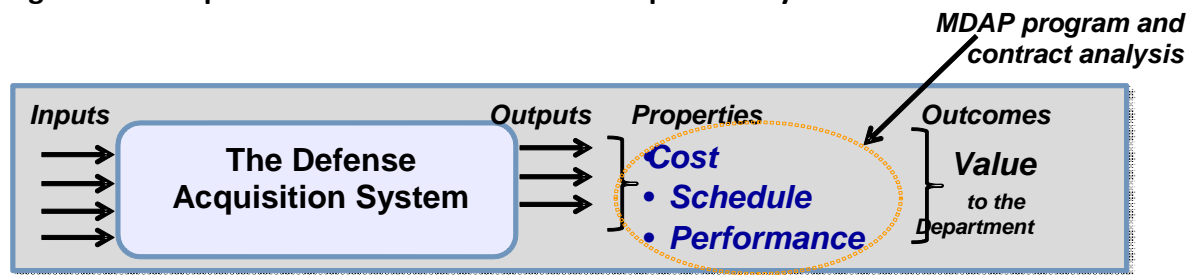
Understanding How Internal Functions and Processes Affect Performance

The acquisition system can be measured at two fundamental levels: (1) the major outputs and outcomes of the system, and (2) the key functions, responsible entities, and institutions accountable within the system to achieve those outputs and outcomes. The most readily available and measurable outcomes assessed throughout the report are cost and schedule growth, but some readily available information on technical performance also is analyzed.

Decomposing the acquisition system into major functional responsibilities enables analysis of how elements of the system affect the ultimate outcomes of interest. Intermediate outputs and outcomes of key institutional functions may correlate with cost, schedule, and performance outcomes, but others may be too small or difficult to discern from available data. Still, a

functional decomposition helps to understand how well the defense acquisition system performs, based on management principles and intermediary outputs and outcomes. As this work moves forward, our greatest challenge will be to identify the relationships between and among the factors the Department can affect (policies, contract terms, incentives, workforce skills, etc.) and the outcomes we are trying to achieve. This report is a step in that process.

Much of our analysis is statistical, focusing on institutional outcomes and their trends rather than on single acquisitions and outliers (see Appendix A for detailed discussion of the statistical methodologies employed). The objective is to see how well we are doing, learn from these generalities, and change our policies and tradecraft as we seek to improve outcomes. Many of the results continue to leverage readily available data on programs and contracts as sets, examining them from different groupings and perspectives. We continue to look for statistically significant differences on sufficient sample sizes to avoid overgeneralizing from case studies.

Scope of Outcomes: Programs or Their Constituent Contracts

Our analyses often examine two main types of performance data:

- **Program-level Data**—describing measurements across the entire program (e.g., estimated final total cost growth from Milestone B (MS B) baseline for all units to be procured), and
- **Contract-level Data**—describing measurements on one of the many contracts that constitute a program (e.g., the total cost growth from original negotiated contract target cost for an early lot of units procured).

Program-level measures show how well the acquisition system developed the ability to produce the overall program against original baselines despite quantity changes, while providing insight into whether cost growth may have been a factor in quantity changes.

Contract-level measures provide early indicators of potential program-level issues by examining performance when the Department contracts for specific work from industry. Nearly all the actual research, development, and production on weapon systems are performed by industry partners through contracts with the Department. Thus, examining performance at the contract level provides detailed and potentially useful indicators of performance that eventually will be seen at the more aggregate program level.

This report often switches between these types of data as we examine different types of institutions (e.g., DoD-wide to military departments to acquisition commands) and different phases of acquisition (e.g., development or early production).

While contracts are the key execution elements of a program (i.e., most goods and even services are provided by contractors), they have different baselines (e.g., contract cost targets) set at different times than the program's MS B baseline. Performance on individual contracts can be measured earlier than their effects might show up in program-level measures. However, because there are often numerous contracts within a program, an individual contract

performance may not necessarily reflect the performance revealed in program-level measurements. **Thus, it is important to recognize what type of data is being discussed at each point in the report.**

Also, care must be taken to note whether cost data have been adjusted for inflation. The available program-level budget data we used have been adjusted for inflation (i.e., reported in "base-year" dollars), but some contract-level cost-growth data have not been adjusted (i.e., are only reported in "then-year" dollars, and insufficient temporal information was available for us to adjust the reported figures for inflation). Thus, partly because of inflation, the program-level cost-growth figures in this report may be lower than those for some contract-level analyses. Cost and price growth in our analysis of margins, however, have been adjusted for inflation.

Measuring Performance on Contracts

Price, schedule, and technical performance are key contract outcomes of interest. Ultimately, the cost to the contractor of providing a good or service relates in various ways to the price paid by the government. Thus, we often examine cost, price, or both (when possible).

Some datasets in this report contain cost, price, and schedule data along with profit or fee (expressed as margin or markup), which allows us to analyze incentive effectiveness. Generally, we were able to adjust these cost data for inflation and thus present cost and price growth in real terms.

Analysis of Work Content Growth and Cost-Over-Target

In other datasets, we do not have profit or fee data but can break down total cost growth into two broad elements of **work content growth** and **cost-over-target**. Work content growth is simply the change in the contract budget base (CBB, which reflects the target cost) since contract initiation. Cost-over-target is the difference between the current CBB and the program manager's estimate at completion (PM EAC). Unless otherwise indicated, all of these contract cost data are reported in "then-year" dollars and are thus *not* adjusted for inflation. Last year's USD(AT&L) report discussed the potential causes of growth in these measures.

Analysis of Contract Margins, Price, Cost, and Schedule

Much of the contract analysis in this year's report involves margin, contract type, and contract competition along with cost, price, and schedule performance for recent MDAP development and procurement contracts. Unfortunately, CBB and PM EAC are not available for these contracts, so we are unable to use the prior approach to assess top-level work content growth and cost-over-target. Most (95 percent) of the contracts had start dates from January 2000 through September 2013. Nearly all were very near completion (i.e., had a DD250 report or completed approximately 95 percent of spend), while a selected few were far enough along to begin seeing if they are having problems (e.g., at least 55 percent of the way through their original schedule). We had 83 development and 83 production contracts that met these criteria in our dataset. Development contracts had about a 70/30 split between cost-plus and fixed-price (by count), whereas production contracts has the reverse (about a 30/70 split).

All *costs* are those to the contractor; *price* is what is paid by the government. Cost to contractor is simply price minus profit or fee. *Margin* equals profit or fee divided by price. Note that this is different than markup, which is profit or fee divided by cost.[5]

Growths (cost, price, and schedule) are simply the final minus initial values all divided by the initial estimated value. In this analysis, we were able to adjust all dollars for inflation (i.e., cost, price, and margin growths are in real terms).

---

[5]*markup = margin / (1 − margin)*; conversely, *margin = markup / (1 + markup)*. Thus, margin and markup are very close when in single percent digits but diverge as they get larger. Arithmetically, margin can approach minus infinity on the low end and has a maximum of 100 percent. Conversely, markup has a minimum of −100 percent and has a maximum that can approach positive infinity.

(This page was left blank intentionally)

2. ACQUISITION SYSTEM OUTCOMES AND TRENDS

A key to improving acquisition is learning from our successes and failures. Without looking at results of past actions, we have no feedback to let us know whether hypotheses and beliefs pan out in the complicated world of defense acquisition. Objectively examining the relative effectiveness of acquisition components and institutions while attempting to distinguish which factors and variables affect outcomes not only allows us to identify successes and failures but also begins to lead us to specific lessons we can try to replicate—and control points we can exploit.

The following analyses examine key outcomes of cost, schedule, and technical performance of MDAPs across the DoD and by Components, commodities, and prime contractors measured at program and contract levels. Combined, these analyses provide insight into cause-and-effect relationships, focusing attention on problems as early as possible, clarifying misunderstandings, and informing assessments and learning.

For our analyses of program data, note that the MDAPs examined are in a varying state of maturity—from early programs that may or may not develop future problems, to mature programs adding new capabilities to existing systems, to completed programs.

For our analyses of contract data, note that each MDAP may have more than one major contract in our datasets. These contracts may be for development or production. Datasets consisted of all readily available cases, including outliers (examining extreme successes and failures are important for learning), rather than a statistical sampling from a larger set. We often report medians because this is a better measure of central tendency for skewed distributions than averages (i.e., arithmetic means), which exaggerate the effect of extreme outliers.[6]

CONTRACT COST GROWTH: POSSIBLE RELATIONSHIP TO ACQUISITION EXECUTIVE DECISIONS

We updated the total cost growth on last year's set of development MDAP contracts in each Military Department shown by who was the USD(AT&L) in office at each contract start date. These data introduce the magnitude and distributions of cost growths in defense acquisition

---

[6]Part of the skewing in the distribution of cost change is the mathematical boundary of cost change because cost cannot decrease more than 100 percent, but it can increase more than 100 percent.

and the challenges of trying to identify causes and motivate performance ownership. A more extensive update of recent contracts will be included in next year's report.

Policy and execution decisions by DoD executives should bear (in part) on the effectiveness of the overall acquisition system during their tenures. Such decisions include changes to the defense acquisition system policies and procedures (e.g., through changes in Departmental regulations); approvals, certifications, and exemptions within that system; institutional organization, policies, and processes; incentives; personnel selection, training, and mentoring; guidance and execution on larger programs, including acquisition strategies and choices; and myriad other effects. More specifically, the acquisition executives chair the boards that review programs at major milestones, guiding both program directions and specific approaches to contracting. Thus, one way to reveal executive effectiveness is to measure cost growth on major contracts started while they were in office.

Tracking effectiveness during the tenure of key executives may help baseline the effectiveness of the defense acquisition system overall and could indicate areas for further research into broad policies and actions leading to improved practices. Note that programs started in the most recent tenures (e.g., Under Secretaries Ashton Carter and Frank Kendall) may have further changes in observed effectiveness due to the relative immaturity of these efforts.

Figure 2-1 shows approximately 20 years of contract data on total cost growth relative to initial contract cost targets for major MDAP development contracts. Figure 2-2 shows the same for early production contracts. Superimposed on the scatter charts are the tenures of Defense Acquisition Executives (DAEs) in place at the time of the contract start date. This was not a statistical analysis of correlation between DAE and contract but an exploratory examination that reveals the wide variation in cost growth of major contracts and possible relationships in time. An assumption was made that generally a contract start date closely follows a key review by the DAE for the programs to proceed with the contract (e.g., the Pre-EMD review or Milestone B decision), although there may be some situations where this was not the case and other factors also can be involved. Outliers and selected contracts for well-known programs are identified by program name (including some duplicates for multiple contracts for the same MDAP). The scatter plot reveals significant variation and skewing in total contract cost growth measured from original contract cost target.

These contracts completed at least about 30 percent of original contract dollar amount (or about at least 55 percent of original schedule). Programs that will exhibit performance problems generally exhibit significant cost growth by this point. However, many recent contracts are not completed and may end with cost growths higher than currently estimated. Thus, there is a data bias of lower cost growth in more recent years that will require more time to determine. For example, total cost growth increased significantly since our prior update of these data on contracts for Global Hawk, GCSS-A, and DDG 1000 (among others) in development, and LCS, F-35 LRIPs, and E-2D in early production.
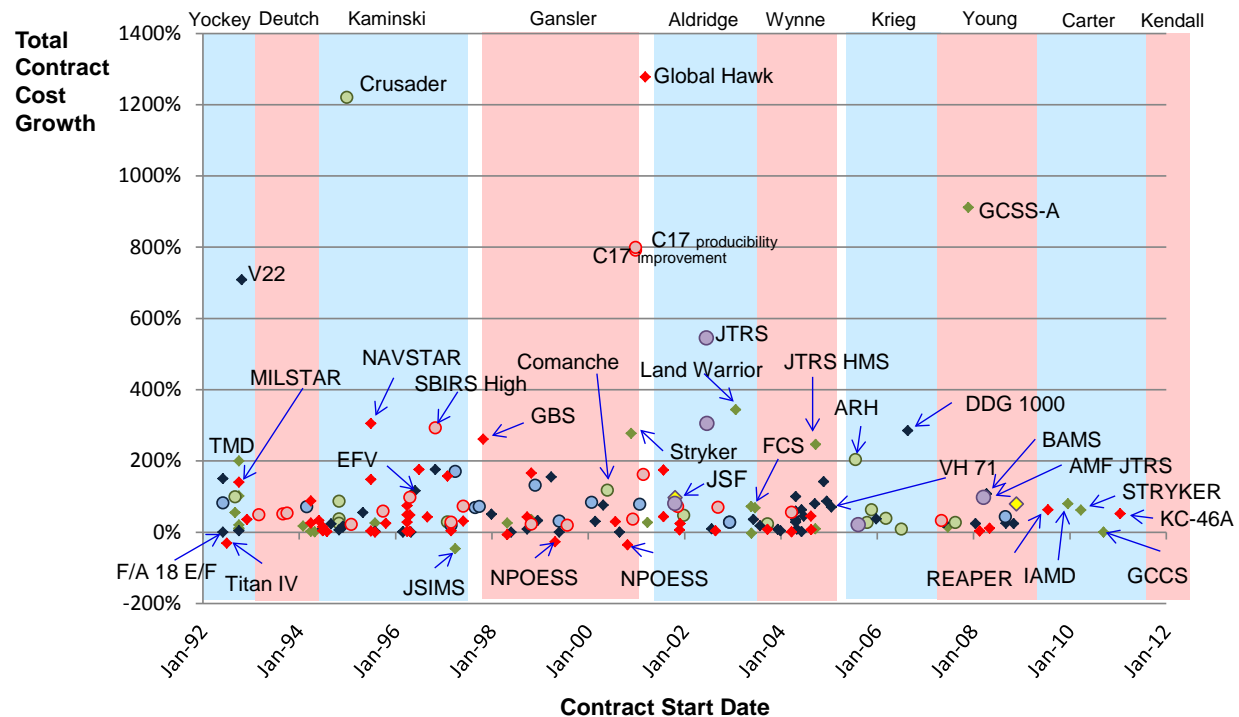
Figure 2-1. DoD-Wide Development Contract Total Cost Growth and USD(AT&L) Tenures (1992–2014).

NOTE: Normally, contract start dates should be relatively close to prior major reviews (usually Milestone B decisions) by the MDA to approve contract award. Army programs are shown in green; Navy in blue, Air Force in red, DoD-wide in purple and yellow. Diamonds on the charts indicate growths mostly attributable to the start of the contract; circles represent significant work added later. Any white bars between Defense Acquisition Executive shaded regions represent periods where there was no confirmed executive. Source cost data were reported in "then-year" dollars (unadjusted for inflation).

Figure 2-2. DoD-Wide Early Production Contract Total Cost Growth and USD(AT&L) Tenures (1993–2014).
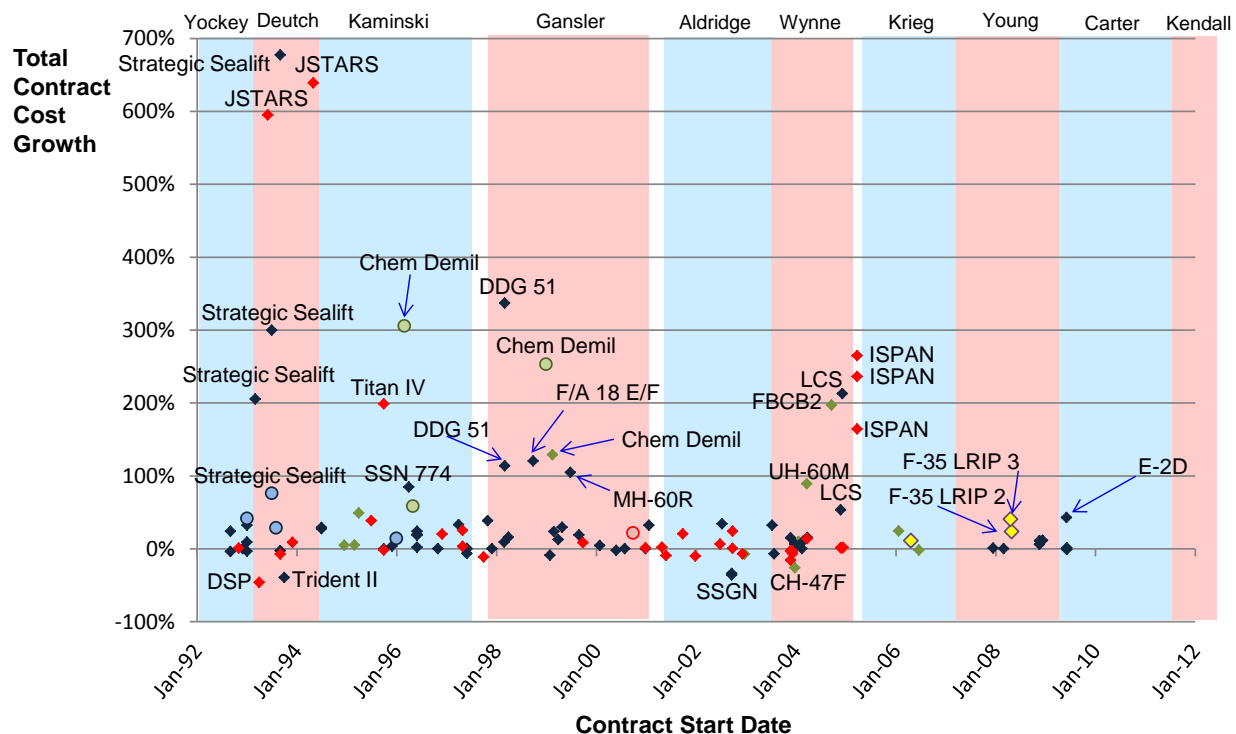
NOTE: Normally, contract start dates should be relatively close to prior major reviews (usually Milestone B decisions) by the MDA to approve contract award Army programs are shown in green; Navy in blue, Air Force in red, DoD-wide in purple and yellow. Diamonds on the charts indicate growths mostly attributable to the start of the contract; circles represent significant work added later. Any white bars between Defense Acquisition Executive shaded regions represent periods where there was no confirmed executive. Source cost data were reported in "then-year" dollars (unadjusted for inflation).

TECHNICAL PERFORMANCE OF MDAPS

While most of this report discusses outcome measures of cost and schedule, this section summarizes some readily available independent assessments of technical performance of weapon systems. Future reports will continue to expand this area.

Mission Effectiveness and Suitability of MDAPs by Organization

One measure of technical performance of acquisition programs is how they rate, as a group, in operational effectiveness and suitability as assessed by the DoD Director of Operational Test and Evaluation (DOT&E).[7] Operational effectiveness is defined in the JCIDS Manual as: "Measure of the overall ability of a system to accomplish a mission when used by representative personnel in the environment planned or expected for operational employment of the system considering organization, doctrine, tactics, supportability, survivability, vulnerability, and threat." Operational suitability is a composite evaluation that considers a system's safety, interoperability, availability, maintainability, and reliability. Operational effectiveness and suitability are not measured solely on the basis of system technical performance parameters. Rather, measurements are accomplished through an evaluation that includes the system under test and all interrelated systems (including weapons, sensors, command and control, and platforms) needed to accomplish a combat mission in expected environments.

Robust developmental testing occurs throughout the earlier phases of a program's life cycle, intended to provide feedback to designers to verify performance and to discover and correct issues so that, by the time operational testing is done on production representative test articles, discovery of major performance issues should be rare.

The following figures summarize DOT&E's assessments of technical performance of weapon systems grouped by Military Department (Figure 2-3) and commodity types (Figure 2-4). The percentage reported represents the number of MDAPs rated Effective or Suitable divided by the number MDAPs assessed. These results were taken from DOT&E's reports prior to any decision to proceed to full-rate production of an MDAP. Each program is rated (or not) as a whole as Effective and Suitable.

Compared to last year's USD(AT&L) report, we changed from temporal line plots of 3-year bins to totals because the sample size in each line plot bin made the plots too sensitive to sample size. Further breakdown to show recent results will be shown in Figure 2-29 and Figure 2-30 in the military department section of the report.

---

[7] DOT&E is independent statutorily from the acquisition organizations and is responsible for, among other things, reporting the operational test results for all MDAPs to the Secretary of Defense, USD(AT&L), Service Secretaries, and Congress.

Figure 2-3. Percent of MDAPs by Military Department Rated as Operationally Effective, and Suitable (1984–2013)
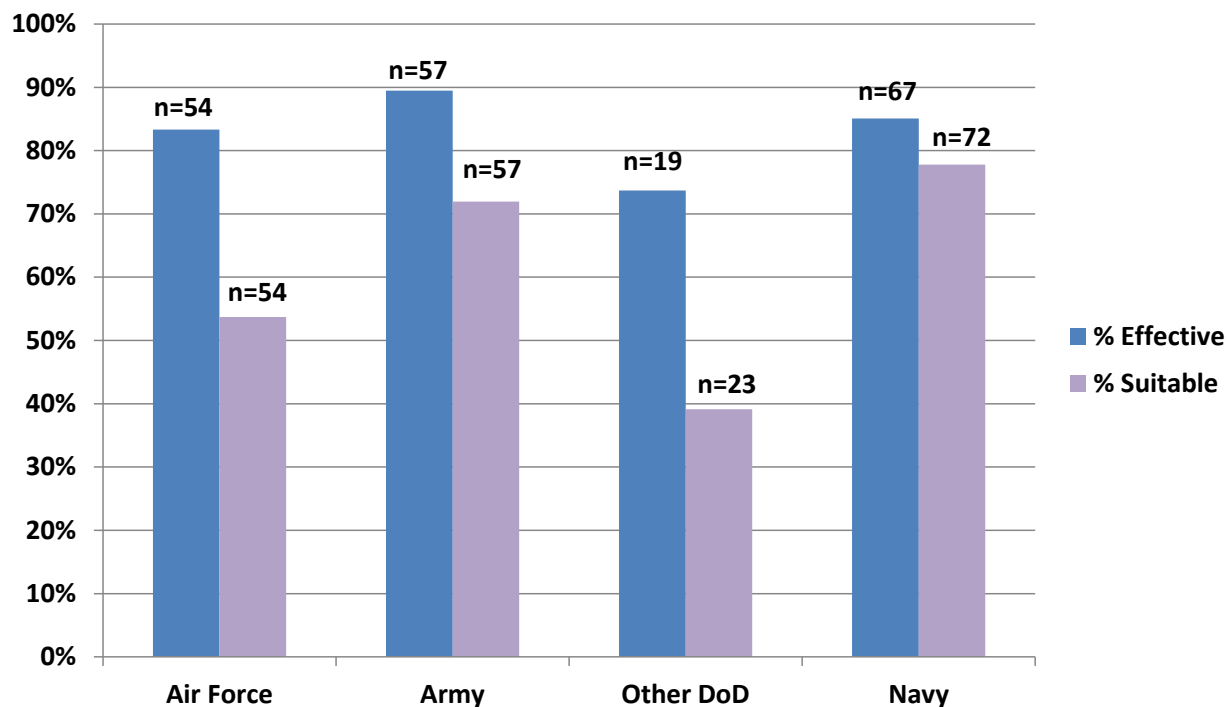
Source: DOT&E reports. NOTE: DoD programs were joint or other programs that are not exclusive to a single Component. Sample sizes differ for some Components because there was not always a definitive binary judgment for effectiveness and suitability in all reports.

Mission Effectiveness and Suitability of MDAPs by Commodity Type

Figure 2-4 shows the aggregate data on operational test results since DOT&E was established in 1984—broken out by the type of systems (also referred to as a commodity within the DoD). Note that the chart includes a mix of programs by dollar value. Breaking out technical performance by commodity reveals domains where additional attention is often needed to improve the operational effectiveness and suitability.

When examining the overall ratings by commodity type, satellites were the most consistent at effectiveness and essentially tied with munitions and ships for suitability. Compared to last year's USD(AT&L) report, C3I and sensors were combined into C4ISR because they work together closely from a mission perspective. Also, UAVs were broken out from aircraft because they are newer and have vertically integrated remote control systems. UAVs exhibited the worst technical performance with only 71 percent of programs being rated *effective* and only 29 percent of programs rated *suitable* in the DOT&E reports. Most UAVs grew out of technology demonstration programs that went quickly into production and fielding, as opposed to out of more standard acquisition programs. This suggests that a lack of engineering discipline and rigorous reliability growth programs is an important factor in these results. Note that some of these categories have very small datasets and thus are less reliable as a basis for generalization.

16

Figure 2-4. Program Ratings in Operational Testing by Commodity Type (1984–2013)
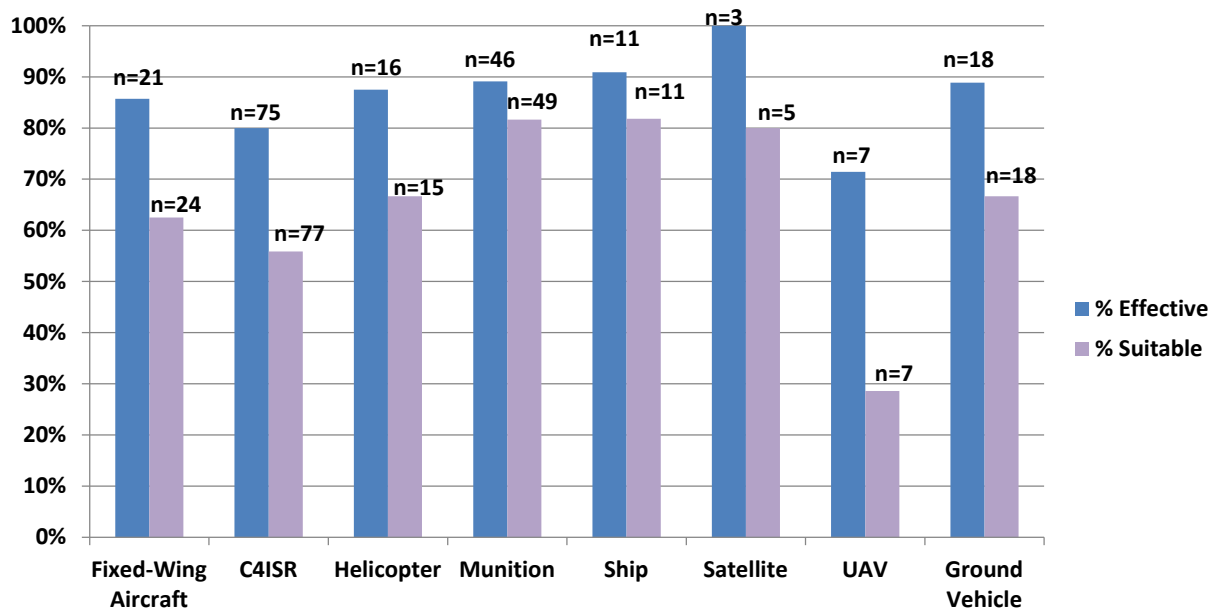
Source: DOT&E reports. NOTE: Sample sizes differ for some Commodities because there was not always a definitive binary judgment for effectiveness and suitability in all reports.

INPUT MEASURES

We now examine two measures of inputs into the defense acquisition system for which the Department has established goals: competition rates and small-business participation.

Rates of Competitive Contracting across All of the DoD for Goods and Services

Competition—or at least creating competitive environments—is a central tenet of our Better Buying Power initiatives. Competition is the single best way to motivate contractors to provide the best value (i.e., the best performance at the lowest price). We have set a strategic objective to increase the percentage of spending on competed contracts from current levels.

Figure 2-5 plots the percentage of all DoD contract dollars that were competitively awarded from FY2006 to FY2013. Trends were established starting in FY2010. Since that year, we have had declining percentages despite our stated goals. Various Better Buying Power initiatives have been established to turn this trend around.

Table 2-1 provided a breakdown for FY2013 of our performance by major Components, and Table 2-2 shows the breakdown by PSC portfolios. While the Air Force and Army met their individual goals, the DoD as a whole did not do so as a result of lower performance by other Components relative to their assigned goals. We will be applying increased management attention to this area going forward. Declining budgets are a factor in reducing opportunities for competition, but the Department will be working hard to reverse this trend.

17

Figure 2-5. Competition Trends: Goals and Actuals (FY2006–FY2013)

Competitively Awarded Obligations
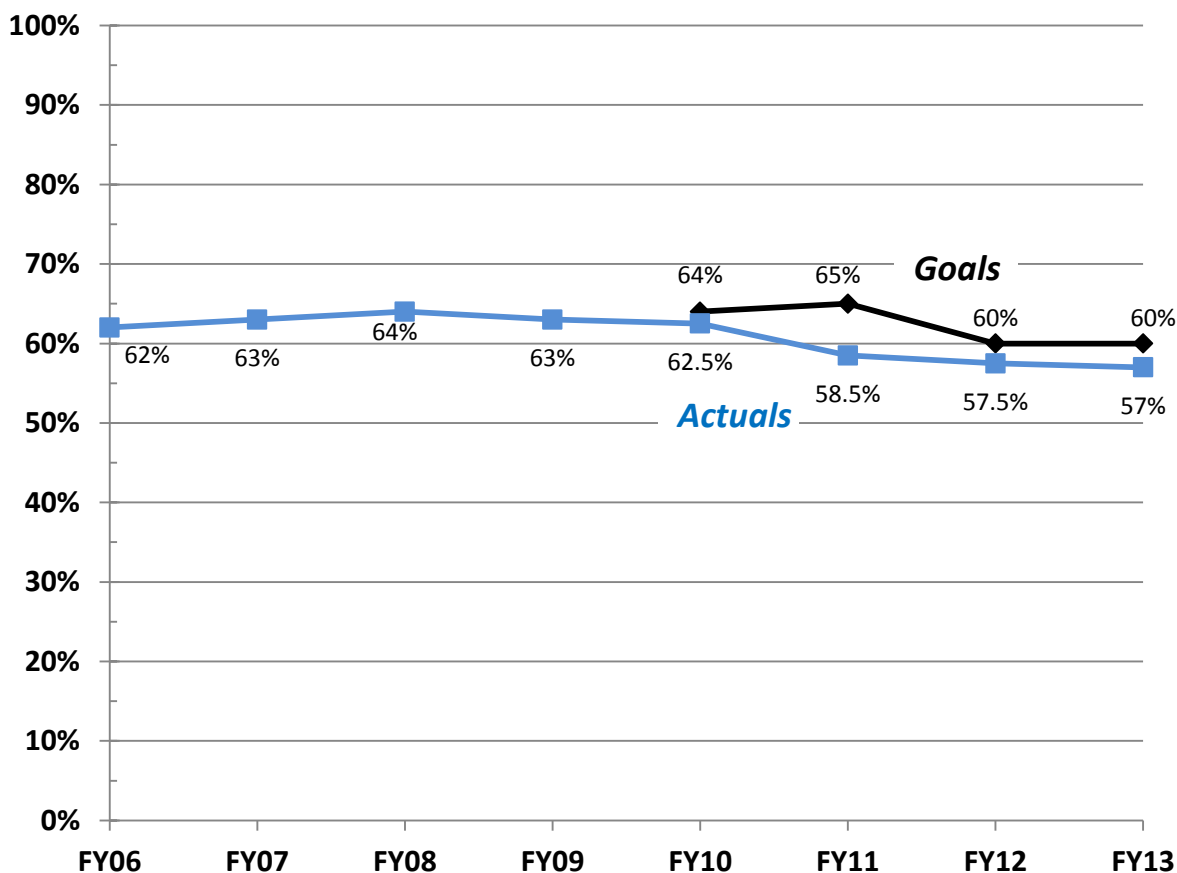
NOTE: We did not establish goals until FY2010. Fraction of contracts competitively awarded are measured on a dollar basis.

Table 2-1. Competition Rate by Obligated Dollars Across Major Components (FY2013)

Competition Rates (by obligations; FY13)	All contracts		Supplies and equipment	Contracted services
	goal	actual	actual	actual
DoD-Wide	60%*	**57%**	39%	73%
Air Force	38%	**41%**	25%	51%
Army	65%	**66%**	39%	80%
Navy and Marine Corps	47%	**41%**	20%	72%
Defense Logistics Agency	86%	**82%**	82%	87%
Other Defense Agencies	81%	**75%**	39%	84%

Table 2-2. Competition Rate by Obligated Dollars Across PSC Portfolios (FY2013)

Supplies and equipment	DoD-wide (actual)	Contracted services	DoD-wide (actual)
Vehicles	16%	R&D	64%
Weapons & Ammo	22%	Equipment	61%
Electronic & Comm.	50%	Electronic & Comm.	66%
Sustainment	56%	Logistics	81%
Textiles & Subsistence	64%	Transportation	86%
Facilities	84%	Facility	76%
Misc.	94%	Construction	90%
		Medical	88%
		Knowledge	69%

NOTES: *The DoD goal for competitive contracts are set in the Performance Improvement chapter of the DoD's annual Overview of the budget request (e.g., USD[Comptroller], 2013, pp. 8–7) and the annual Strategic Management Plan (e.g., DCMO, 2013). Component goals are set internally by the DoD.

Small-Business Participation

We pursue small-business utilization goals as required by law, but more importantly, because of potential benefits from small-business contributions. More small-business engagement can increase the competitive nature of our solicitations, resulting in better cost and schedule performance on contracts. Small businesses can also infuse new technical solutions as we pursue capabilities for our warfighters.

Figure 2-6 shows actual Department-wide small-business utilization (obligations) relative to yearly goals. Recent trends since FY2011 have been improving and we have nearly reached our FY2009 level. However, we have not achieved our goal and remain below FY2003 to FY2005 levels.

Small-business eligible dollars obligated to small businesses in FY2013 totaled $47.2 billion across the Department: $15.9 billion for products (i.e., supplies and equipment) and $31.3 billion for services (see Table 2-3). Overall, we missed our FY2013 goal of 22.5 percent by 1.3 percentage points. Table 2-4 shows small-business obligation rates by portfolio groups.

Figure 2-6. Small-Business Utilization Trends: Goals and Actuals (FY2001–FY2013)
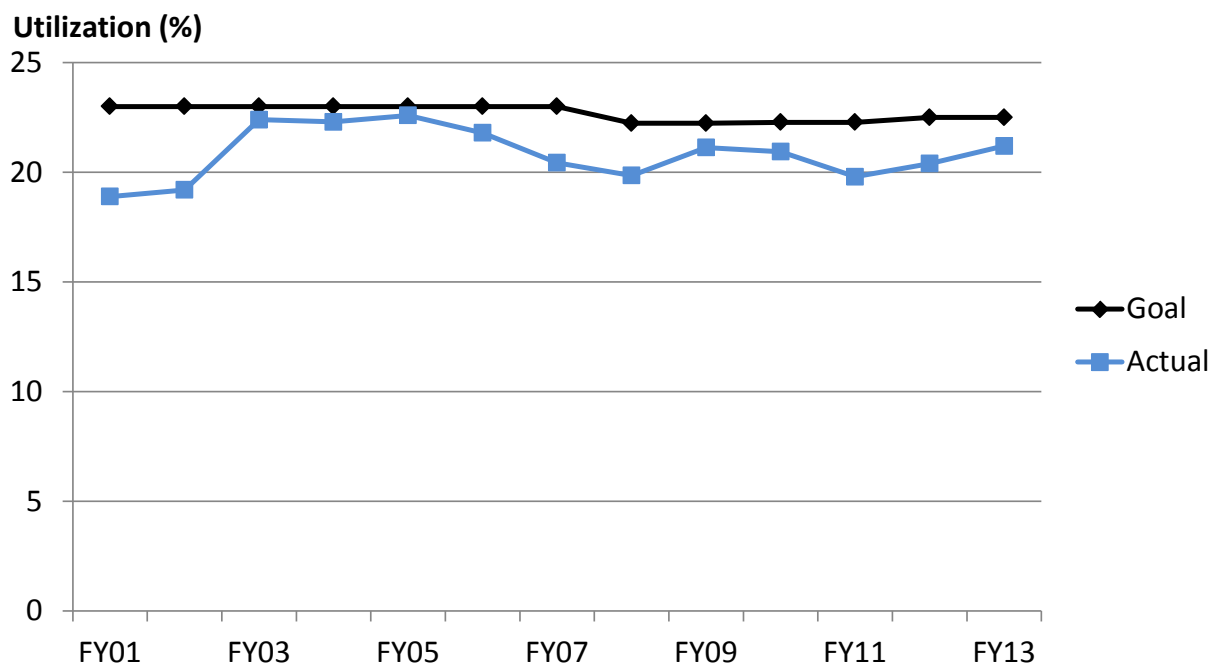
Table 2-3. Small-Business Obligations, DoD-Wide and by Components (FY2013, by obligated dollars)

Small Business Obligations	FY13 Goal	Total		Products		Services	
		(TY$,B)	(%)	(TY$,B)	(%)	(TY$,B)	(%)
Army		$ 17.4	27.4%	$ 3.8	17%	$ 13.6	33%
Navy & Marines		$ 11.8	15.2%	$ 2.7	6%	$ 9.1	27%
Air Force		$ 6.4	14.5%	$ 1.1	8%	$ 5.2	18%
Defense Logistics Agency		$ 7.4	37.6%	$ 6.8	38%	$ 0.6	36%
Other Defense Agencies		$ 4.3	24.1%	$ 1.5	38%	$ 2.7	20%
Total	**22.5%**	**$ 47.2**	**21.2%**	**$ 15.9**	**15.4%**	**$ 31.3**	**26.2%**

NOTES: Percentage of small business obligations is on a dollar basis. Eligible dollars exclude categories of actions as specified by law (see http://www.acq.osd.mil/osbp/gov/goalingExclusions.shtml). The threshold for what qualifies as a small business is based on the North American Industry Classification System (NAICS) code applicable to the contract. Budget amounts are in billions of unadjusted, then-year (TY) dollars.

Table 2-4. Small-Business Obligations by Portfolio Group, DoD-Wide (FY2013, by obligated dollars)

PRODUCTS	Small Business Obligations (TY$,B)	(%)
Vehicles	$ 0.3	1%
Weapons & Ammo	$ 0.6	6%
Electronic & Comm.	$ 4.7	28%
Sustainment	$ 5.9	25%
Facilities	$ 2.5	45%
Textiles & Subsistence	$ 2.0	27%
Misc.	$ 0.0	1%
TOTAL	**$ 15.9**	**15%**

SERVICES	Small Business Obligations (TY$,B)	(%)
R&D	$ 3.7	14%
Equipment	$ 1.7	12%
Electronic & Comm.	$ 3.6	25%
Logistics	$ 0.5	15%
Transportation	$ 0.7	20%
Facility	$ 8.1	48%
Construction	$ 3.8	41%
Medical	$ 1.1	65%
Knowledge	$ 7.9	26%
TOTAL	**$ 31.3**	**26%**

NOTES: Percentage of small business obligations is on an obligated dollar basis. Eligible dollars exclude categories of actions as specified by law (see http://www.acq.osd.mil/osbp/gov/goalingExclusions.shtml). The threshold for what qualifies as a small business is based on the NAICS code applicable to the contract. "Ammo" is ammunition; "Comm." is communication. Budget amounts are in billions of unadjusted (TY) dollars.

COST AND SCHEDULE PERFORMANCE: OVERALL

Nunn-McCurdy Program Breaches

Each MDAP is required by law to submit a Selected Acquisition Report (SAR) to the Congress 45 days after the President's annual budget submission and under various other circumstances (see: 10 U.S.C. section 2432; Schwartz, 2010). The SAR reflects what is included in the President's Budget as well as a comprehensive summary of MDAP cost, schedule, and technical performance measures. Historical SAR data serve as the primary sources for much of the program-level analysis in the report due to their relative availability and comprehensiveness.

Common cost measures such as Program Acquisition Unit Cost[8] (PAUC), which includes both RDT&E and procurement, and Average Procurement Unit Cost[9] (APUC), which includes only procurement), are codified in statute. The statute also requires that programs exceeding certain thresholds (measured by PAUC or APUC changes relative to their original and latest

---

[8] Section 2432(a)(1), Title 10, U.S.C. defines program acquisition unit cost as "the amount equal to (A) the total cost for development and procurement of, and system-specific military construction for, the acquisition program, divided by (B) the number of fully configured end items to be produced for the acquisition program."

[9] Section 2432(a)(2), Title 10, U.S.C. defines procurement unit cost as "the amount equal to (A) the total of all funds programmed to be available for obligation for procurement for the program, divided by (B) the number of fully configured end items to be procured."

program baselines) must go through a rigorous reexamination and certification to Congress along a variety of specified criteria. This process commonly is referred to as the "Nunn-McCurdy" process, named for the two original sponsors of the legislation.

Two types of breaches are called out in the Nunn-McCurdy process: *significant* and *critical*. A "significant" breach is the lower threshold and is intended to warn Congress that a program is experiencing high unit-cost growth. A "critical" breach signifies the cost growth is even higher, triggering the formal reexamination and certification process mentioned above. The criteria for a significant breach are 15 percent from the current baseline reported in the previous SAR or 30 percent cost growth in APUC or PAUC from the original baseline. A critical breach occurs when the program experiences 25 percent from the current baseline or 50 percent cost growth from the original baseline. Normally, the event of breaching (crossing a threshold) is counted once for each threshold as a means for measuring program performance even though technically we track programs that are in breach throughout the remainder of its life.

Figure 2-7 shows the number of Nunn-McCurdy breaches since 1997, including each individual breach for those programs that have breached multiple times (e.g., a significant breach followed by a critical breach of the original baseline). The NDAA for FY2006 made changes to the Nunn-McCurdy statute, adding the requirement to report unit-cost growth from the original baseline; this additional requirement caused the large spike shown in 2005, where 11 programs had to report preexisting significant breaches. There have been 90 total breaches since 1997 and 43 over the last 8 years (not counting the adjustment year immediately following the statutory change). The average number of total breaches starting in 2006 has been about 5.4 per year. Thus, the number of breaches in 2013 is just below average but much higher than in 2012. The average number of critical breaches starting in 2006 is about 3.9 per year—higher than in 2013. Still, it is too early to tell if we are on a downward trend. Also, some breaches are caused partially or totally by quantity changes, particularly as budgets decrease. In future reports we will look more closely at the root causes of this year's Nunn McCurdy breaches using PARCA analysis, expanding on results from last year's report.

Figure 2-7. Nunn-McCurdy Breaches per SAR Year (1997–2013)
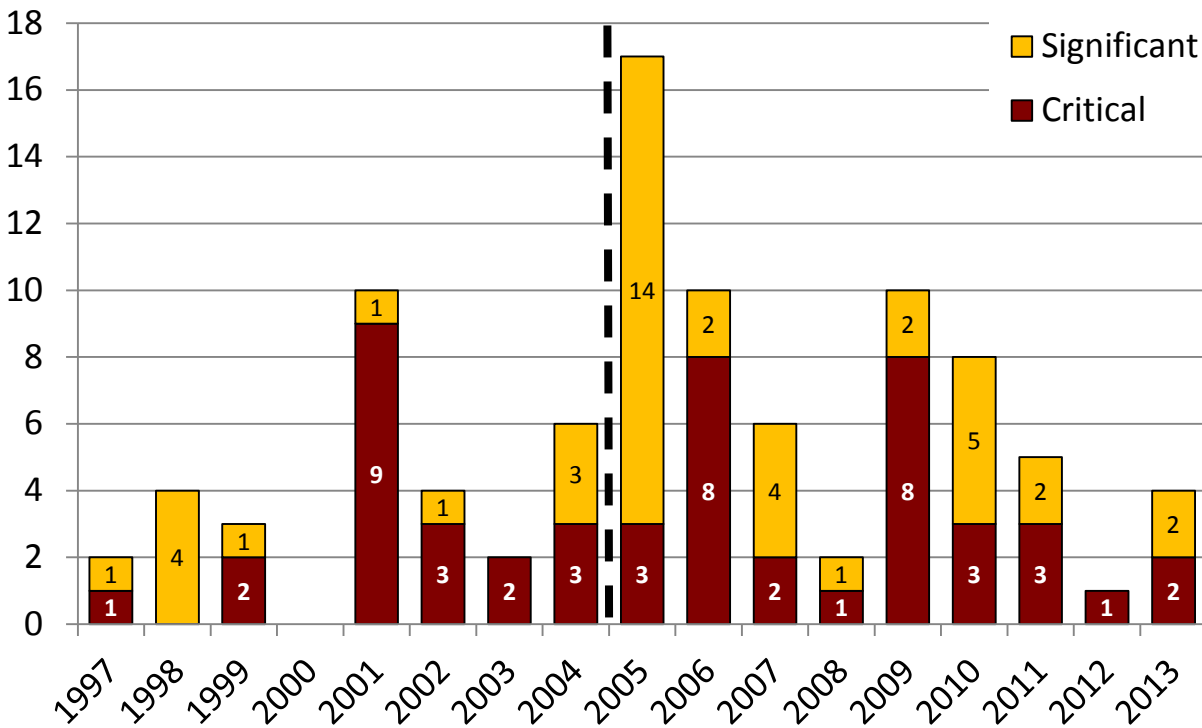
NOTE: The criteria for breaches were changed in NDAA 2006, affecting counts starting with 2005. Breaches are determined using "base-year" dollars (i.e., adjusting for inflation). This plot includes the number of breaches in each annual SAR reporting cycle, which nominally equates to calendar year but may include updates early in the following calendar year from the President's Budget Request. Breaches in different years for different thresholds or baselines for the same program are included in each respective year. If a program reported both a significant and critical breach in the same year, it is only plotted here as a critical breach. Nunn-McCurdy breach reporting was established in the NDAA for FY1982, so the count shown here for 1997 may differ from that by others, depending on whether prior notification for the same breach without rebaselining has occurred.

To show how MDAP costs compare against baselines at any one point, Figure 2-8 lists the MDAPs in both PB14 and PB15 along with their percent APUC growth from original baseline in those budgets. In this particular example, about half of the programs are above their original baselines and half below (in other cases with other measures and baselines, the percent above can be worse or better). As expected, program-level cost growth can change from year to year due to internal program performance, external budgetary changes, or other factors. Some programs improved their APUC significantly between PB14 and PB15 while some remained essentially the same, and still others became significantly worse.

The Department's cost estimates (upon which acquisition baselines are based) assert it is about equally likely that the actual costs will be higher or lower than the estimate (assuming we acquire the same system as estimated). This assertion appears to be borne out by Figure 2-8 in that about half of the programs are above their baseline and half below. Of course, this is just one snapshot, and others do show more cost growth than cost reduction relative to baseline.

Also, while some programs exhibit extremely poor cost control and breach the Nunn-McCurdy thresholds, others may be executing well and even below their baselines.

Table 2-5 below summarizes a different analysis of Nunn-McCurdy breaches by commodity. In this case, we do not "double count" programs that have breached multiple times. This allows us to compare the types of programs that have poor cost performance (as evidenced by crossing any Nunn-McCurdy threshold) to those that have never breached during this period.

Thirty-two percent of all MDAPs since 1997 have had either a significant or critical breach (1 percentage point higher than last year's number). This analysis appears to show that all commodities are susceptible to breaches, with the chemical weapons demilitarization (Chem Demil), space-launch (Evolved Expendable Launch Vehicle—EELV), and helicopter programs having the highest breach rates (note EELV is the only space-launch program in this dataset). Chem Demil and EELV are unique in their missions and management schemes, but it is unclear at this point why helicopters breach at a higher rate than other standard commodity types.

The Nunn-McCurdy process provides insights but only comes into play when programs already are having problems. Also as indicated above, even though they are unit metrics, PAUC and APUC are sensitive to quantity changes, which can create the appearance of or mask cost growth due to acquisition planning and execution problems. Explaining cost growth as simply a result of changing quantity, therefore, can be complicated and misleading without careful analysis. The SARs do contain cost variance discussions that provide some useful explanations of changes in individual programs, but inconsistencies in the variance concepts and process make it difficult to understand root causes of cost changes, especially across multiple programs and from the quantitative data in the cost variance reports (see, for example, Hough, 1992; Arena et al., 2006). Recent efforts by the Department—especially the root-cause analyses summarized in last year's report—are aimed at getting beyond mere characterization of "symptoms" and proximate causes to gain understanding of the underlying root causes and mechanisms leading to cost growth on programs.

Figure 2-8. MDAP APUC From Original Baseline (PB14 and PB15)

*Original
Baseline*

JPALS Inc 1A — 96%
VTUAV — 71%
SSN 774
H-1 Upgrades
MH-60S
AGM-88E AARGM
AH-64E Remanufacture
PAC-3
NAS
WIN-T Inc 2
LPD 17
AWACS Blk 40/45 Upgrade
Tactical Tomahawk
GMLRS/GMLRS AW
IDECM/IDECM Blocks 2/3
MH-60R
CH-47F
CH-53K
GPS III
UH-60M Black Hawk
RMS
SM-6
JTRS HMS
Trident II Missile
DDG 51
IDECM/IDECM Block 4
Excalibur
AMRAAM
C-130J
RQ-4A/B Global Hawk
LHA 6
PIM
F-35/F-35 Engine
KC-130J
E-2D AHE
MQ-4C Triton
B61 Mod 12 LEP TKA
B-2 EHF Inc 1
EA-18G
JLTV
V-22
MQ-9 Reaper
F-22 Inc 3.2B Mod
SSC
F-35/F-35 Aircraft
DDG 1000
G/ATOR
Patriot/MEADS CAP/Missile
HC/MC-130 Recap
KC-46A
AMDR
AEHF/AEHF SV 1-4
FAB-T
P-8A
CEC
C-5 RERP
WGS
CVN 78/EMALS
CVN 78/CVN 78
MUOS
EELV
SBIRS High/Baseline (GEO 1-4,…
NMT
AH-64E New Build
JASSM/JASSM-ER
MQ-1C Gray Eagle
JASSM/JASSM Baseline
SBIRS High/Block Buy (GEO 5-6)
AIM-9X Blk II
JDAM
AEHF/AEHF SV 5-6
WIN-T Inc 3
SDB II
GBS
IAMD
JSOW
AMF JTRS   < -50%
MIDS

| PB14
♦ PB15

Percent Change from APB
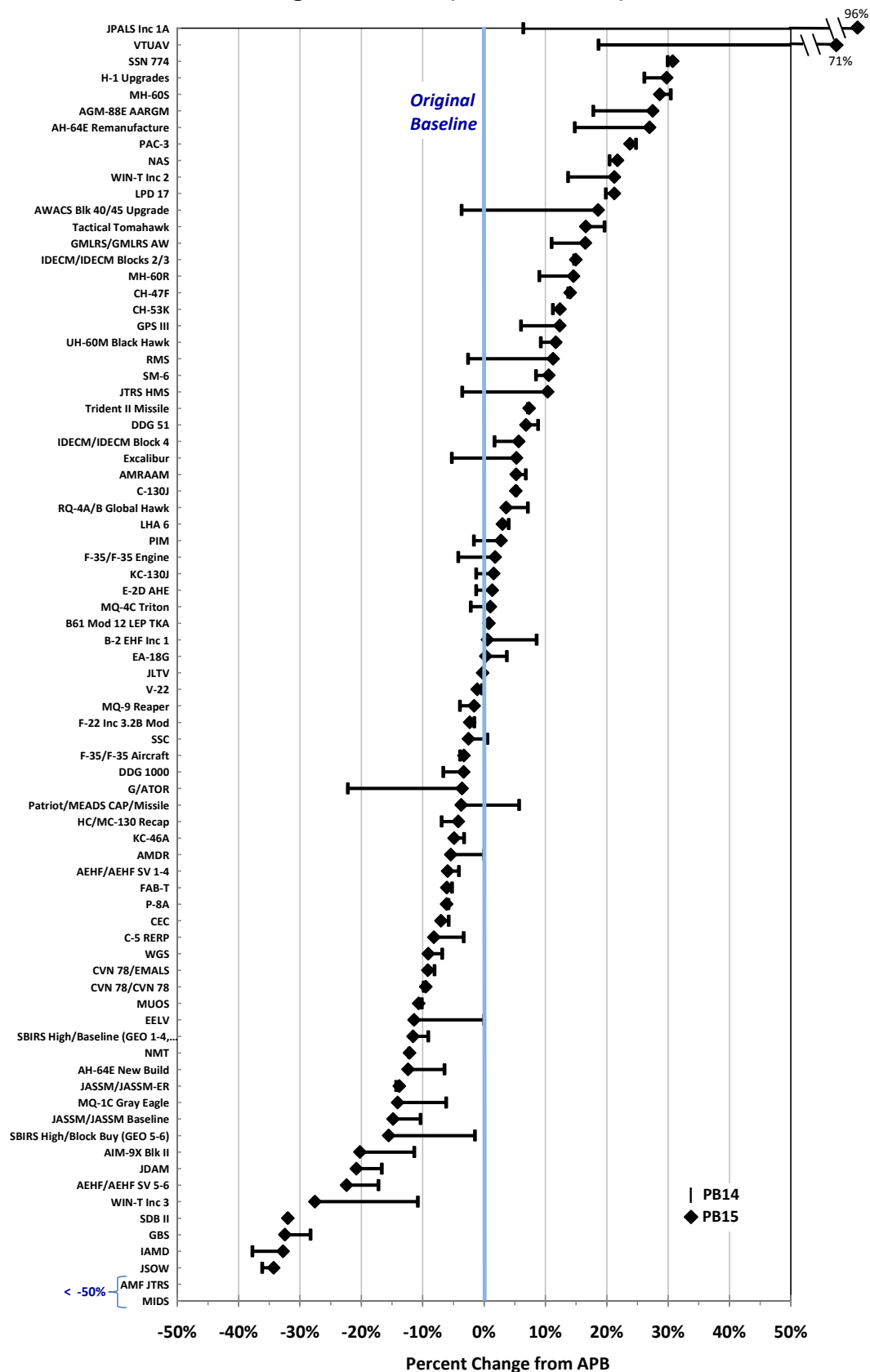-50% -40% -30% -20% -10% 0% 10% 20% 30% 40% 50%

25

Table 2-5. Fraction of MDAPs by Commodity Type that Crossed Any Nunn-McCurdy Threshold (1997–2013)

Commodity Type	Total # of Programs	# of Programs that Ever Breached	Breach Rate	# of Programs with at Most a Significant Breach	# of Programs with at least one Critical Breach
Chem Demil	4	4	100%	1	3
Space launch	1	1	100%	—	1
Helicopter	13	8	62%	4	4
Satellite	12	5	42%	2	3
Fixed-Wing Aircraft	27	11	41%	4	7
UAV	6	2	33%	—	2
Ground Vehicle	11	3	27%	2	1
Ship	19	5	26%	1	4
Munition/Missile	29	7	24%	2	5
C4ISR	51	11	22%	4	7
Missile Defense	8	1	13%	—	1
Total	**181**	**58**	**32%**	**20**	**38**

NOTE: Compares number of programs that have crossed any Nunn-McCurdy threshold to those that have never crossed a threshold. Breaches are determined using "base-year" dollars (are adjusted for inflation). These commodity types are slightly different from those reported last year. For example, sensors logically belong in the C4ISR category, and UAVs are broken out from aircraft to help reveal how they have fared.

MAIS and Business Systems

An area of particular performance concern is the performance of Major Automated Information Systems (MAIS), including business systems. MAIS are Acquisition Category IA (ACAT IA) programs that either meet a specified budgetary threshold[10] or are designated by the Secretary of Defense (or designee) as a MAIS (e.g., due to risk or other concerns) (see 10 U.S.C., section 2445a). Title 10 mandates various reports and baselining mechanisms for MAIS. Our performance at acquiring these systems has been uneven at best, and understanding the reasons for our poor performance by MAIS programs is a high priority.

Figure 2-9 shows the current dollar size of the MAIS programs reported in at least one MAIS Annual Reports (MARs) from 2011–2013.[11] At least half of the MAIS programs have original

---

[10]For example, one MAIS threshold is $378 million in FY2000 dollars for total life-cycle costs (see 10 U.S.C. section. 2445 for details and other thresholds).

[11]MARs are the MAIS equivalent of SARs and are provided to Congress to satisfy the requirement in 10 U.S.C., section 2445b.

total cost baselines below about $500 million (except in the Navy) while others can cost billions of dollars.
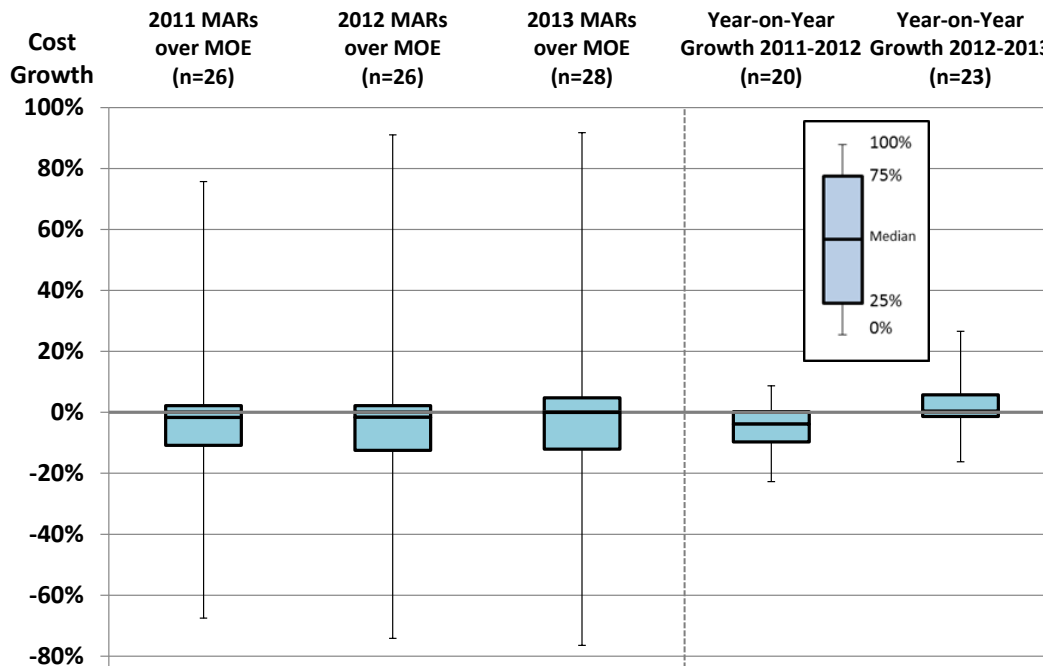
Figure 2-10 shows the cost growth from original baselines as reported in the 2011–2013 MARs. Cost growth at the medians has been at or just below zero. Figure 2-11 compares MAIS cost growth from original baselines based on whether USD(AT&L) or designee is the MDA (i.e., "IAM" programs) or whether oversight has been delegated to the Component Acquisition Executive (i.e., for "IAC" programs). Here we find that most of IAC programs have negative cost growth while IAM programs have medians at or just above zero.

This preliminary analysis supports anecdotal assertions that, unlike MDAPs, MAIS may be changing their scope to available funds and schedule thresholds—especially at the Component level. We currently do not have centralized data to measure requirement changes across MAIS to test this hypothesis. If true, however, then examinations of cost performance relative to baselines would not be as meaningful as they are on MDAPs (where general capabilities are assumed to be relatively stable).

Figure 2-9. Size Distributions (by Dollars) of MAIS Programs (2011–2013 MARs)
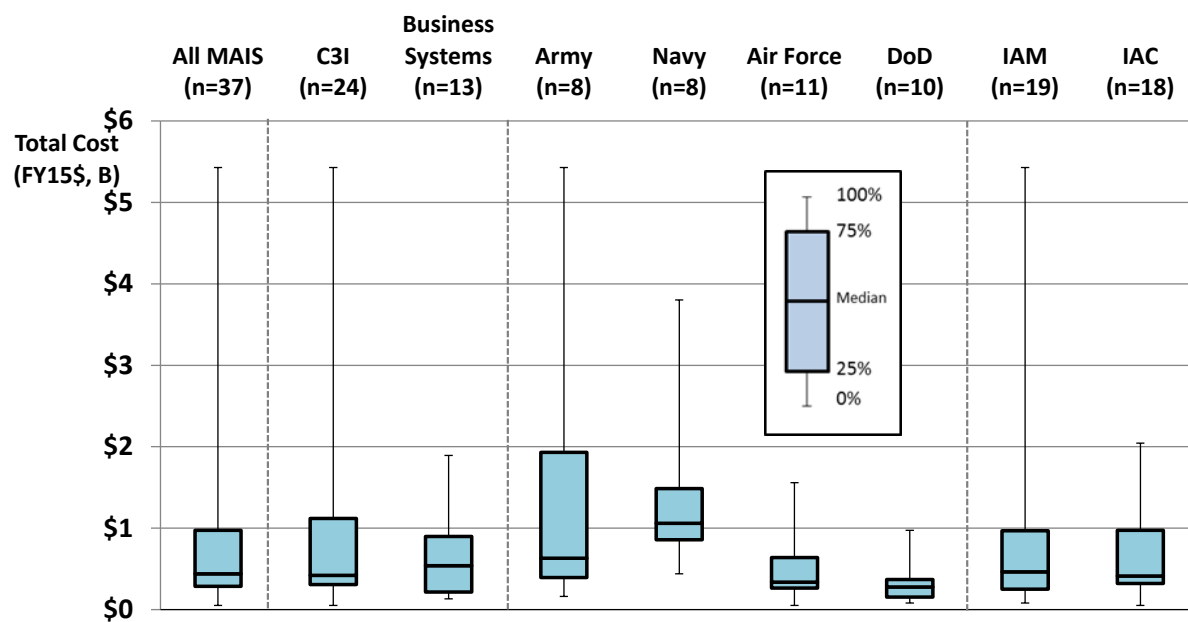
NOTE: Costs are total cost estimates at original baseline as reported in the 2011–2013 MARs in billions of FY2015 adjusted dollars (FY15$, B). Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Figure 2-10. MAIS Acquisition Cost Growth From Original Baseline and Yearly (2011–2013)

NOTE: "Year-on-Year" shows cost growth between successive MARs. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Figure 2-11. MAIS IAM and IAC Total Cost Growth From Original Estimate (2011–2013)
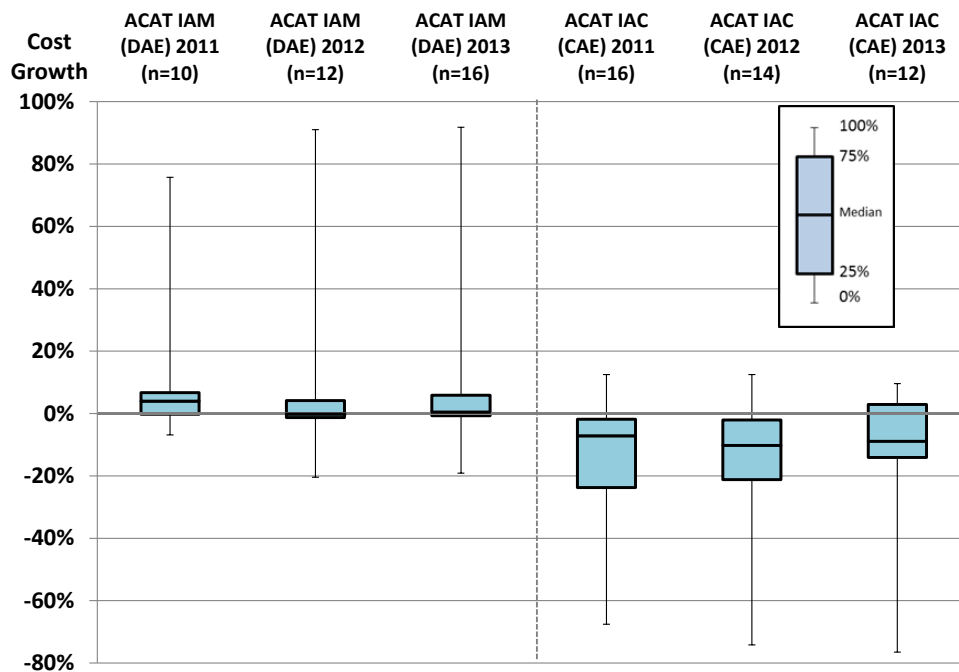
NOTE: Total cost includes O&M and Working Capital Fund costs. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

With respect to schedule growth, the MAIS reporting in the 2011–2013 MARs exhibited a median growth of about 2 to 3 months with three-quarters of the MAIS at or below 6 to 10 months (see Figure 2-12). These include all reporting MAIS, so there is concern that schedule growth can increase significantly before the programs are completed.

Figure 2-12. MAIS FDD Schedule Growth From Original Estimates (2011–2013 MARs)
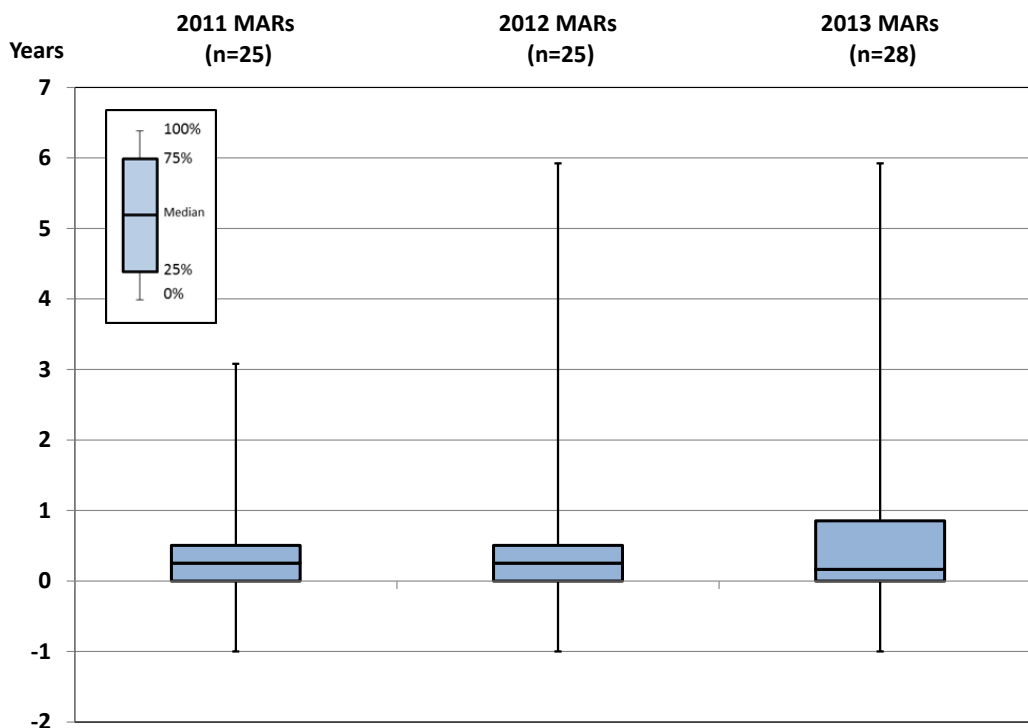
NOTE: Original estimates are those reported in the first MAR for each MAIS. Schedule period is from MS B or FFO to FDD. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

In addition to these analyses, note that any information systems designated as MDAPs (instead of MAIS) are included in our analysis of Nunn-McCurdy cost breaches earlier in this report.

Operation and Support Costs

Life-cycle costs after production of MDAPs are often as large as or larger than their RDT&E and procurement costs. Thus, understanding how our acquisition system performs relative to O&S has been an important objective of our performance analysis.

O&S measures are affected by such factors as reliability, availability, and maintainability that relate (at least in part) to system design and quality. Unfortunately, current O&S data measures are affected by numerous factors that change over time independent of system design aspects such as operational assumptions and the number of people required to operate a system. Changes in operational environments outside design specifications can also affect reliability, availability, and maintainability, resulting in changes to actual O&S costs.

The SARs contain total lifetime and annual unit O&S costs for the MDAPs. Historically, these estimates have been somewhat inconsistent. For example, nothing prevented a program office from reporting unit O&S cost estimates on a per item basis (e.g., per airplane) in one year and on a per usage basis (e.g., per flying hour) in the next year. Beginning with the December 2012 SARs, USD(AT&L) now requires more consistent use of units for calculating O&S estimates, but it will take years for enough data to be available to examine O&S cost-growth trends.

Despite these reporting changes, the underlying assumptions about system use and reliability have—and still can be—changed from year to year without affecting raw O&S cost estimate numbers. For example, the dollar value of consecutive O&S estimates might be identical, but one estimate assumed 100 flying hours per year and a mean time between failure of 25 hours, while the next estimate assumed 90 flying hours per year and a mean time between failure of 30 hours (both of which would serve to reduce annual O&S costs). In other words, important program O&S cost-performance issues such as underlying reliability may not be revealed by simply monitoring top-level O&S measures between SARs.

Also, DoD Components have not consistently tracked actual O&S costs against final estimates after the systems are fielded. Consequently, estimates may lack important data on historical actuals.

Despite these deficiencies in reporting, a sufficient number of O&S reports exist in the SARs to generally assess recent O&S cost estimate growth. Since the late 1990s:

- MDAPs' unweighted average growth in *total* O&S cost estimates has been about 8 percent per year in real terms; and
- MDAPs' unweighted average growth in per *unit* O&S cost estimates has been about 4 percent per year in real terms.

These data point to a significant factor in the upward trend of O&M and its dominance in DoD's budget noted earlier (see Figure 1-2), particularly if the SARs generally underestimate O&S costs.

COST AND SCHEDULE PERFORMANCE: DEVELOPMENT

First, we examine cost growth for MDAPs in development. For this year's report we used additional data sources to expand our insight into contract performance, especially into how incentives tie to performance outcomes. Some data are updated, and next year's report will provide further updates.

Recent Temporal Patterns of Contract Cost Growth: Development

In the 2013 report, we identified seven common patterns of cost growth using Earned Value data. Such characterizations can be leading indicators of programs that may be in, or heading for, trouble. Increases in the CBB show work content being added to the contract, potentially

indicating premature starts or work added later. Increases in the PM EAC potentially indicate poor estimating and/or poor management, depending on the divergence from the CBB.

Table 2-6 summarizes our characterization of cost-growth patterns in MDAP development contracts for programs that had a MS B approved (original or new) after January 21, 2009. We included contracts that started near or after this date. Many other recent contracts did not have enough earned value trend data to be included in this analysis, but eight did. Generally, these few contracts had about the same ratio of poor to good cost control as the 176 MDAP contracts with start dates from 1993 to 2011 discussed in last year's report. Of those eight contracts with sufficient data, one-fourth exhibited patterns indicative of well-managed programs that may not incur significant cost growth by completion. This percentage is about the same as the 28 percent we found last year. The six contracts with a significant cost-growth pattern varied in their temporal patterns.

This sample of eight recent contracts is too small for precise comparisons, but it indicates that cost growth continues to require the oversight attention we are giving them to control cost. USD(AT&L) regularly reviews execution on MDAPs and their major contracts, including the six contracts in the table with significant cost growth patterns. In one case, the contract type is an FPIF where the cost has passed the point of total assumption, which caps the Department's exposure to pay for continued cost growth. In others, we continue to look for ways to reduce costs through revised incentive structures, efficiency investments, and control of engineering requirements and changes.

Table 2-6. Cost-Growth Patterns on Recent Development Contracts (Since August 2008)

Dominant Cost Growth in Pattern	Number of Contracts
Cost-over-target *(negligible work content growth)*	2
Work-content growth in first year *(negligible cost-over-target)*	1
Work-content growth in first year followed by cost-over-target growth	1
Work-content growth in first year followed by steady work-content growth	1
Work-content growth after first year followed by steady work-content growth	1
Stable *(no significant changes in work content or cost-over-target)*	2
Total	**8**

Note: Cost growths (work content or cost-over-target) are identified if they are greater than about 15 percent. All contracts had start dates in or after August 2008.

Cost and Price Comparisons on Development Contracts

The question of how cost and price behave relative to each other on the same contract is indicative of the effectiveness of our financial incentives, cost estimating, and contract negotiations. Figure 2-13 compares the relative change of cost and price on 83 prime development contracts since about the year 2000. Overall, cost and price moved in the same direction 94 percent of the time. Both fixed-price (FP) and cost-plus and hybrid (CP/H) contracts

had high instances of cost growth: 90 percent of FP contracts and 82 percent of CP/H contracts had some cost growth. In most cases when costs increased the price increased less, indicating that incentives were working in the right direction.

Two troubling cases are apparent. In some cases, price went up more than costs—most often on CP/H contracts (see the rightmost bar on the plot). Of even more concern, in a very few CP/H cases, price went *up* but costs when *down*; this is unacceptable.

Note that in this data we were not able to control for work content growth on these contracts.

Figure 2-13. Relative Changes in Contract Cost and Price on Prime MDAP Development Contracts (2000–2012)
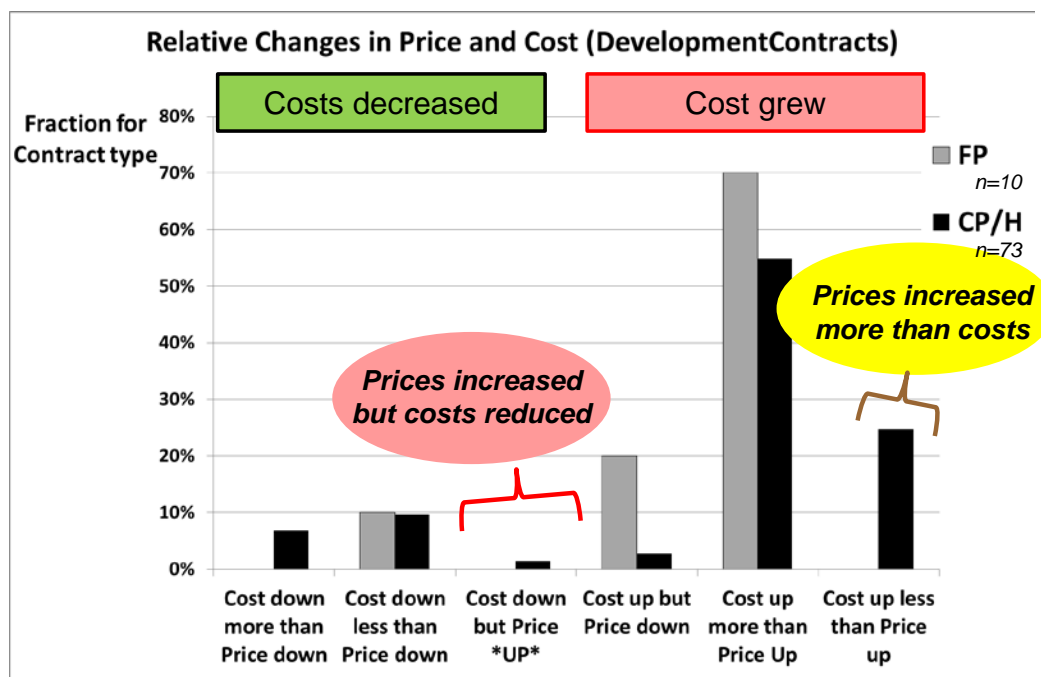
Contract-Level Development Price and Schedule Growth by Commodity

A comparison of recent development contract performance contract-by-contract (i.e., not weighted by cost) revealed differences in price and schedule growth by commodity type for prime MDAP contracts since about the year 2000. Table 2-7 shows that UAVs had the highest price growth since contract award, followed by ships. Missiles had the lowest, with a median price reduction. Schedule growth was the worst for ships, followed by rotary-wing contracts. Ground vehicles and electronics systems had the lowest (negligible) schedule growth (the latter when unweighted by spend).

Table 2-7. Median Price and Schedule Performance by Commodity: Development Contracts (2000–2012)

Medians	Overall	Fixed-wing (n=10)	Rotary-wing (n=9)	UAVs (n=5)	Ships (n=6)	Ground Vehicles (n=8)	Missiles (n=4)	Space (n=8)	Electronics (n=31)
Unweighted (contract basis)									
Price growth	**16%**	17%	23% (v)	Highest * **94%**	40% (v)	7%	Lowest * **−4%**	21%	13% (v)
Schedule growth	**0.4 yr**	1 yr (v)	2.2 yrs (v)	0.4 yr	Highest * **3 yrs** (v)	Lowest * **0.0 yr**	1 yr	1.3 yrs (v)	Lowest * **0.1 yr**
Weighted by spend (dollar basis)									
Price growth	**29%**	20%	25% (v)	Highest *59% (v)	Highest * **78%** (v)	17%	Lowest * **−8%**	38% (v)	5%
Schedule growth	**2.5 yrs**	2.2 yr	Highest *3.8 yrs (v)	Low * **0.0 yr**	Highest * **3.9 yrs**	Low * **0.1 yr**	1.6 yr	2.4 yrs (v)	1.7 yrs

NOTE: * indicates statistically significant compared to at least most others; (v) indicates noticeably higher variability in the distribution. Price growth is adjusted for inflation and is similar to cost growth. Munitions had insufficient sample size to be included in this table. Sample size was insufficient to report on munitions contracts.

Program-Level Cost Growth: Development

Generally, RDT&E costs must be paid regardless of how many units are produced. In that sense, they are a fixed cost for the Department to arrive at a point where we can actually procure and field a capability. Thus, for RDT&E, cost growth could be tracked in total rather than by unit produced to avoid confusing the effects of quantity changes with RDT&E cost growth.

The following figures show total RDT&E cost growth by MDAP portfolio relative to the original program baseline and at 2-year intervals. These different views are useful because they show how the portfolio performed from inception compared to how it performed in its most recent periods. Differences that are statistically significant are indicated with asterisks.

Examining RDT&E cost growth from each program's original baseline estimate is important to capture the overall growth since inception; however, it may not be the best choice for gaining insight into recent cost-growth management because MDAPs can have very long lives. When we analyze a program from inception, we are forced to carry all cost growth indefinitely. Programs that are currently executing well and that had a one-time cost increase in the distant past can appear to be poor performers in the long term. Therefore, it is important that we look at both types of data.

Notably, the data show considerable (and sometimes seemingly conflicting) differences between the medians and the arithmetic means. This is because the data are highly skewed, and a single but very large outlier can have a large effect on the mean while not affecting the median.[12] Thus, we show the medians to provide a better sense of the central tendency of the population, and we provide the means for completeness and to show the effects of these outliers. Also, these values are not weighted by program dollar value, so they reflect program effectiveness generally regardless of size.

For each analysis, we first show the main portion of the cost-growth distribution (between −10 percent and 100 percent growth) followed by a separate figure showing all outliers (especially those with growth greater than 100 percent). To be consistent with the other plots in this year's report, the "box and whisker" charts on each plot show quartiles (last year's report showed 20th and 80th percentiles). Medians are the lines within each box. Gray-shaded columns in the table beneath each chart were periods with very low sample counts because full SAR reporting was not made in those years due to new Presidential administrations. The "x" markers above the box mark the five largest overruns (although outliers above 100 percent only appear on the outlier graphs).

Figure 2-14 shows that RDT&E total cost growth has been statistically flat since 2005. Growth in 2001 was lower than in 2002–2004, and it increased after that. Thus, the medians since 2010 are not significantly higher than even the 9 percent in 2006. This emphasizes the importance of statistically examining the population rather than just considering the median or mean.
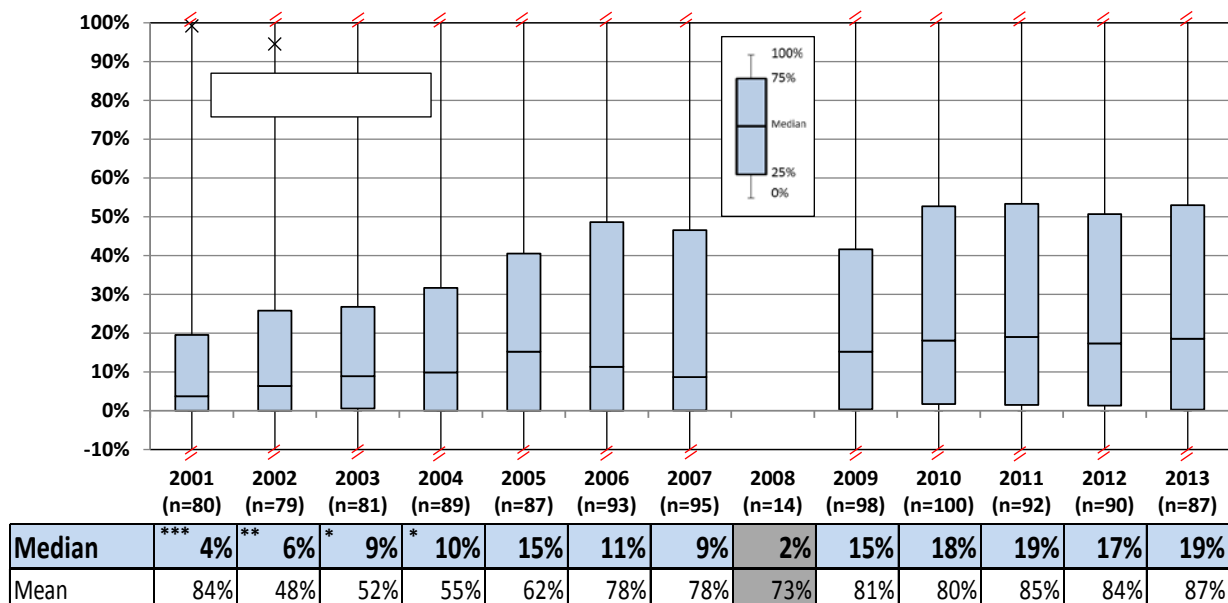
Figure 2-15 plots cost growth on an expanded axis to show all outliers greater than zero. The maximum cost-growth percentages are very high due to a small number of outliers and are not statistically representative of the overall MDAP portfolio. These extreme growths are not due to measurement error and so were not excluded from the analysis. Still, they do skew the aggregate data, which is an important fact for knowing how to measure and discuss cost growth across a program population. Similar skewing is observed in various complex commercial projects (see, for example, Flyvbjerg et al., 2002).

Understanding *why* a program may exhibit such a large percentage increase in RDT&E cost requires an individual examination of each case. For example, in Figure 2-15, the C-130J remains the highest outlier since 2002. This program originally was envisioned as a nondevelopmental aircraft acquisition with a negligible RDT&E effort planned. Several years into the program, a decision was made to install the Global Air Traffic Management system, adding several hundred million dollars to development and causing the total development cost growth to climb upward of 2,000 percent. This is an example of a major change in the program rather than poor planning or execution, although significant program changes like this are not necessarily the reason for all extreme cases of cost growth.

---

[12] Part of the skewing in the distribution of cost change is the mathematical boundary of cost change because cost cannot decrease more than 100 percent but can increase more than 100 percent.

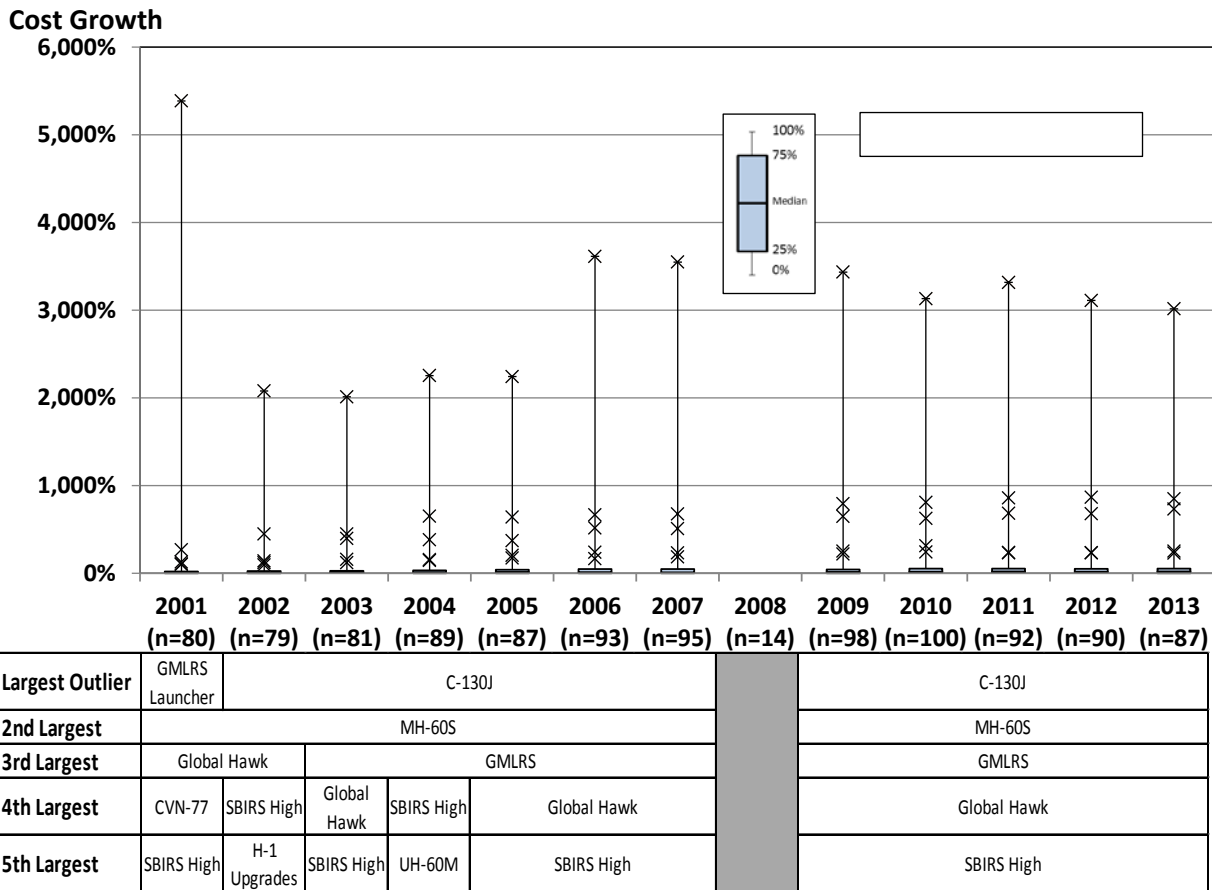**Figure 2-14. RDT&E Program Cumulative Cost Growth Over Original MS B Baseline: Central Quartiles (2001–2013)**

Year	2001 (n=80)	2002 (n=79)	2003 (n=81)	2004 (n=89)	2005 (n=87)	2006 (n=93)	2007 (n=95)	2008 (n=14)	2009 (n=98)	2010 (n=100)	2011 (n=92)	2012 (n=90)	2013 (n=87)
Median	***4%	**6%	*9%	*10%	15%	11%	9%	2%	15%	18%	19%	17%	19%
Mean	84%	48%	52%	55%	62%	78%	78%	73%	81%	80%	85%	84%	87%

*** Statisically lower than 2005-2013   ** Statistically lower than 2009-2013   * Statistically lower than 2010, 2011, 2013

NOTE: Source data were reported in "base-year" dollars (i.e., are adjusted for inflation). Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.
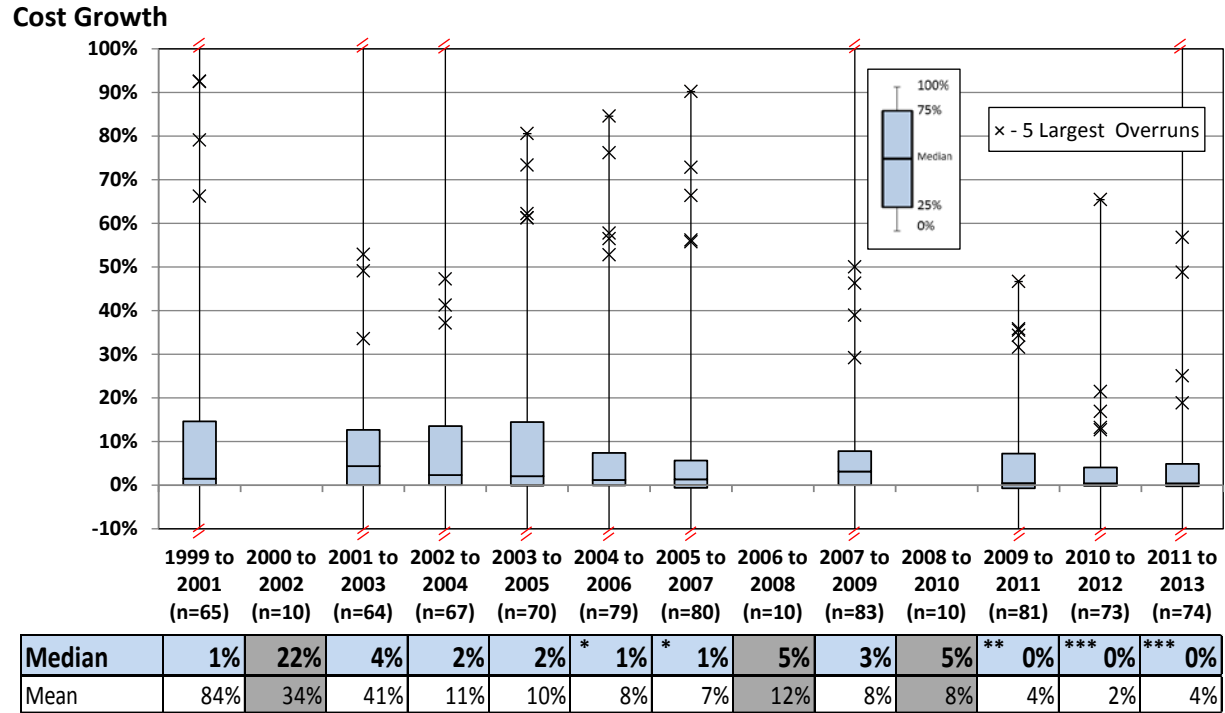
Figure 2-15. RDT&E Program Cumulative Cost Growth Over Original MS B Baseline: Outliers (2001–2013)

	2001 (n=80)	2002 (n=79)	2003 (n=81)	2004 (n=89)	2005 (n=87)	2006 (n=93)	2007 (n=95)	2008 (n=14)	2009 (n=98)	2010 (n=100)	2011 (n=92)	2012 (n=90)	2013 (n=87)
Largest Outlier	GMLRS Launcher	C-130J							C-130J				
2nd Largest	MH-60S								MH-60S				
3rd Largest	Global Hawk		GMLRS						GMLRS				
4th Largest	CVN-77	SBIRS High	Global Hawk	SBIRS High	Global Hawk				Global Hawk				
5th Largest	SBIRS High	H-1 Upgrades	SBIRS High	UH-60M	SBIRS High				SBIRS High				

NOTE: Source data were reported in "base-year" dollars (i.e., are adjusted for inflation).
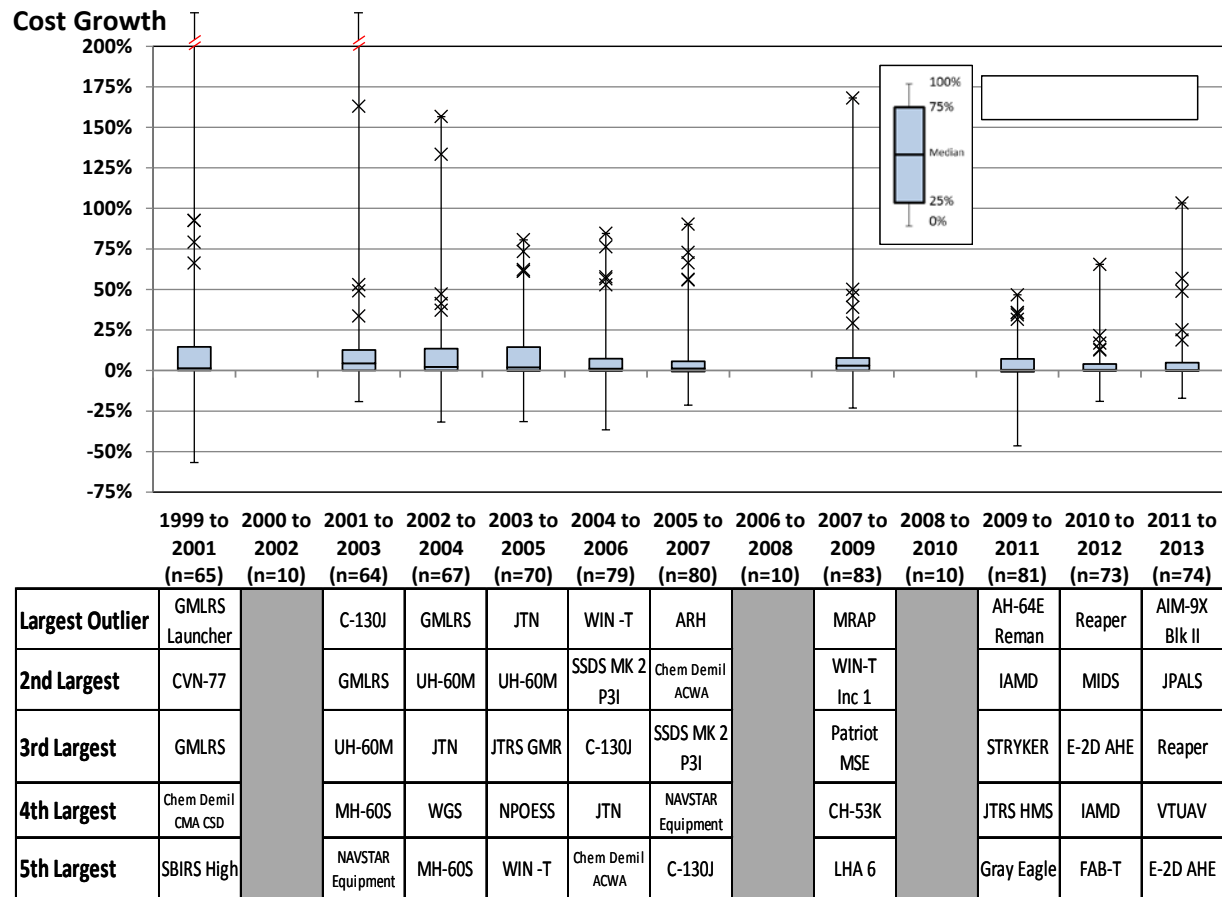
Figure 2-16 shows a continuing improvement trend when comparing program RDT&E cost growth on a 2-year basis. The 2009–2011 period showed statistically lower growth than the 2001–2003 and 2002-2004 periods. The last two periods show further reductions and are significantly lower than periods ending 2001–2005 and the period ending in 2009. With medians at zero percent, they still are insufficient to counter prior RDT&E growth since original baselines. Figure 2-17 shows the five largest cost-growth outliers in these 2-year periods.

Figure 2-16. RDT&E Program Cost Growth Over 2-Year Spans: Central Quartiles (2001–2013)

Cost Growth

	1999 to 2001 (n=65)	2000 to 2002 (n=10)	2001 to 2003 (n=64)	2002 to 2004 (n=67)	2003 to 2005 (n=70)	2004 to 2006 (n=79)	2005 to 2007 (n=80)	2006 to 2008 (n=10)	2007 to 2009 (n=83)	2008 to 2010 (n=10)	2009 to 2011 (n=81)	2010 to 2012 (n=73)	2011 to 2013 (n=74)
Median	1%	22%	4%	2%	2% *	1% *	1%	5%	3%	5%	0% **	0% ***	0% ***
Mean	84%	34%	41%	11%	10%	8%	7%	12%	8%	8%	4%	2%	4%

* Statisically lower than period ending 2003   ** Statistically lower than periods ending 2003-2004 *** Statistically lower than periods ending 2001-2005 and 2009

NOTE: Source data were reported in "base-year" dollars (i.e., are adjusted for inflation). Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Figure 2-17. RDT&E Program Cost Growth Over 2-Year Spans: Outliers (2001–2013)

	1999 to 2001 (n=65)	2000 to 2002 (n=10)	2001 to 2003 (n=64)	2002 to 2004 (n=67)	2003 to 2005 (n=70)	2004 to 2006 (n=79)	2005 to 2007 (n=80)	2006 to 2008 (n=10)	2007 to 2009 (n=83)	2008 to 2010 (n=10)	2009 to 2011 (n=81)	2010 to 2012 (n=73)	2011 to 2013 (n=74)
Largest Outlier	GMLRS Launcher		C-130J	GMLRS	JTN	WIN -T	ARH		MRAP		AH-64E Reman	Reaper	AIM-9X Blk II
2nd Largest	CVN-77		GMLRS	UH-60M	UH-60M	SSDS MK 2 P3I	Chem Demil ACWA		WIN-T Inc 1		IAMD	MIDS	JPALS
3rd Largest	GMLRS		UH-60M	JTN	JTRS GMR	C-130J	SSDS MK 2 P3I		Patriot MSE		STRYKER	E-2D AHE	Reaper
4th Largest	Chem Demil CMA CSD		MH-60S	WGS	NPOESS	JTN	NAVSTAR Equipment		CH-53K		JTRS HMS	IAMD	VTUAV
5th Largest	SBIRS High		NAVSTAR Equipment	MH-60S	WIN -T	Chem Demil ACWA	C-130J		LHA 6		Gray Eagle	FAB-T	E-2D AHE

NOTE: Source data were reported in "base-year" dollars (i.e., are adjusted for inflation). Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

MAIS CYCLE TIMES

In addition to the analysis in last year's report on cycle times for MDAPs, Figure 2-18 adds results on cycle time for the MAIS programs reporting in the 2011–2013 MARs.   The plot shows the number of years for each MAIS from MS B (or the date of "funds first obligated" [FFO]) to the Full-Deployment Decision (FDD) in the original estimate (i.e., the first MAR for the MAIS). MAIS with MS B or FFO before 2009 had a median cycle time of 5 years; since then, the estimated median cycle times dropped to just below 3 years.[13] In other words, before 2009, half of the MAIS were planned with cycle times longer than 5 years. Since 2009, that estimate

---

[13]Many MAIS increments have a MS B but not MS A, so we have more consistent data using MS B. For comparison, 5 years since MS A or FFO (not MS B as shown here) to FDD is the statutory threshold beyond which a certification of variance is required. The end points of the 5-year period have changed over the years, but it is currently from MS A or Preferred Alternative Decision (PAD) to FDD.

has dropped significantly, and no program is planned on taking longer than 4.75 years since MS B or FFO. This appears to be a direct result of the legal requirement for Critical Change Reports if the five year period is breached. Whether the Department achieves these estimates and whether this improves acquisition performance or not has yet to be determined. The median schedule growth on all currently reporting MAIS since their original estimate is about 2 to 3 months (see Figure 2-12).

The optimal cycle time cannot be predetermined absent information on the system in question. In some cases, long cycle times may be a concern given the pace of information technology advancement (often expressed on the order of Moore's Law and correlates of about 18 months). On the other hand, setting arbitrary schedule deadlines may incentivize undesirable management decisions and short cuts, causing failures to meet needs from end users and possibly increasing costs for subsequent upgrades.

Figure 2-18. MAIS Originally Estimated Cycle Time From MS B or FFO to FDD (2011–2013 MARs)
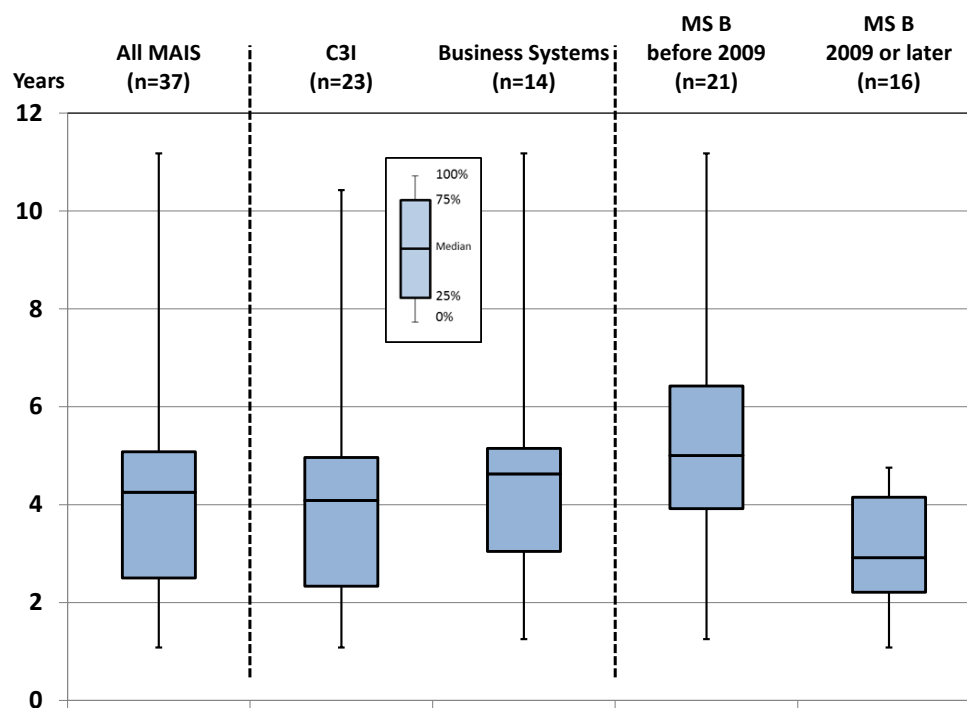
NOTE: Original estimates are those in the MAIS' first MAR. Included are the latest data on programs that appeared in at least one MAR from 2011 through 2013. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

COST AND SCHEDULE PERFORMANCE: PRODUCTION

This section focuses on cost growth for MDAPs in production. As with development, adding 1 year of data to the analysis in last year's USD(AT&L) report of performance since 1970 is

insufficient to have statistical significance, so this year we focused on recent results and insights from a different dataset. Next year we will provide an update on last year's results.

Recent Temporal Patterns of Contract Cost Growth: Production

As with development contracts above, we examined recent production contracts for programs that had a MS B approved (original or new) after January 21, 2009 that had sufficient earned value temporal data to identify its pattern.

However, only one contract met these criteria. It exhibited significant work-content growth in first year followed by cost-over-target growth. While this pattern was not common in the historical dataset discussed in our report last year, a single contract is insufficient to draw conclusions about trends on programs with recent MS B approvals.

Cost and Price Comparisons on Production Contracts

As with development, we examined how cost and price behave relative to each other on the same production contract. Figure 2-19 compares the relative change of cost and price on 83 prime production contracts (including both LRIP and FRP) since about the year 2000. Overall, cost and price moved in same direction 86 percent of the time (a bit lower than in development). Cost growth was less dominant than in development; half of fixed-price incentive firm (FPIF) and CP/H contracts and 30 percent of firm-fixed-price (FFP) contracts had cost growth greater than zero. This difference is due to the dominance of FPIF in LRIP and FFP in FRP.

Again, two troubling cases are apparent. As in development, price went up more than costs most often on CP/H contracts (see the rightmost bar on the plot), and this happens more often than in development for all contract types. In a few CP/H cases, price went *up* but costs went *down*, and this also happened more often than in development for all contract types. These inverse relationships between cost and profit are not acceptable.

However, note that again we were not able to control for work content growth on these contracts.

Figure 2-19. Relative Changes in Contract Cost and Price on MDAP Production Contracts (2000–2012)
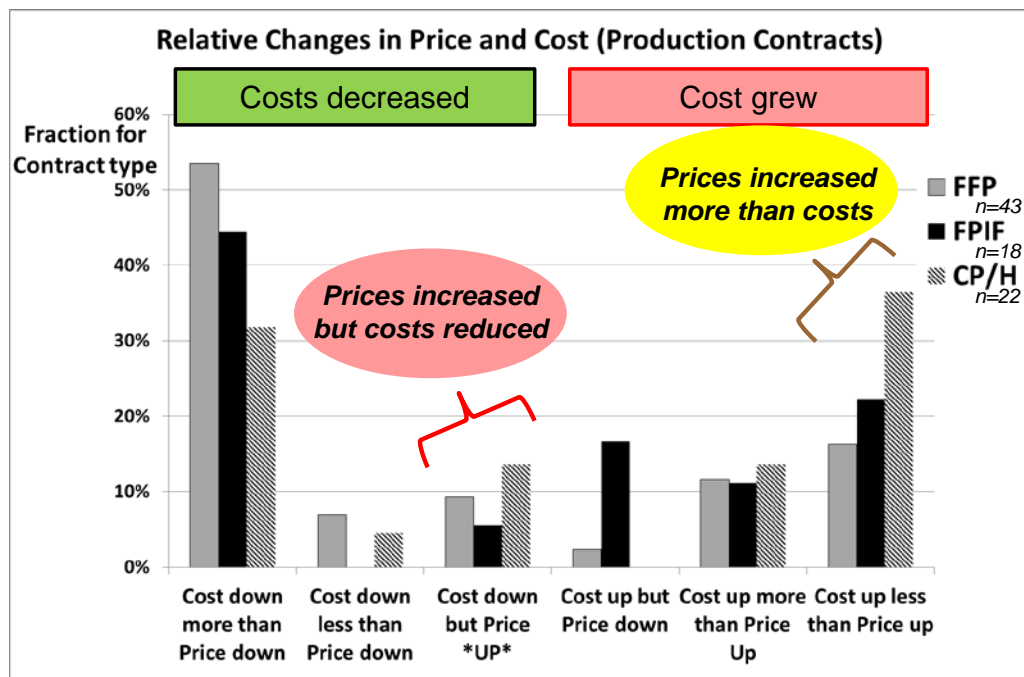
Contract-Level Production Price and Schedule Growth

As we found with development contracts, some performance differences were revealed when comparing recent contract performance by commodity type for prime MDAP contracts since about the year 2000. Table 2-8 shows that contracts for ground vehicles and munitions had the highest price growth on a per contract basis, but when adjusting for spend (i.e., on a per dollar basis), space contracts had the highest price growth, followed by ground vehicles. Further examination revealed that quantity increases and schedule reductions appear to be the major causes of price growth on ground vehicle contracts. These indicators are important, but we have to also examine work-content changes and schedule incentives when reviewing price and cost growth.

Cost growths were similar to price growths, except for missiles where cost growth was somewhat lower.

We also were able to examine final margins on these contracts. Ground vehicles and missile contracts had the highest, but that was a result of the dominance of FFP contracts for these commodity types. Analysis discussed below in Chapter 3 will show that FFP contracts generally provide the highest final margins in our dataset. The missile contracts had good price control and even better cost control associated with their higher margins.

41

Table 2-8. Median Price and Schedule Performance by Commodity: Production Contracts (2000–2012)

Medians	Overall	UAVs (n=9)	Ships (n=24)	Ground Vehicles (n=8)	Munitions (n=4)	Missiles (n=19)	Space (n=13)
Unweighted (contract basis)							
Price growth	−1%	−3%	−2% (v)	Highest * 13% (v)	Highest * 4%	−1%	−2%
Schedule growth	0.6 yrs	Highest * 2 yrs (v)	0.3 yrs	0.0 yrs (v)	0.6 yrs (v)	0.5 yrs	1 yr
Weighted by spend (dollar basis)							
Price growth	12%	Lowest * −4%	0% (v)	Higher * 13% (v)	3%	Lowest * −2%	Highest * 23% (v)
Schedule growth	0.5 yr	Highest * 2 yrs	0.1 yr (v)	0.0 yr (v)	0.2 yr (v)	0.5 yr (v)	Higher * 1.4 yr

*statistically significant; (v) = variable.
NOTE: Median cost growth for missiles was lower than the median price growth, otherwise they were similar. Cost and price growth are adjusted for inflation. Fixed-wing, rotary-wing, and electronics contracts had insufficient sample size to be included in this table.

Program-Level Early Production Cost Growth (Quantity-Adjusted)

Now at the program level, the following figures summarize the unit procurement cost growth across the MDAP portfolio from the original MS B baseline and in 2-year increments. Again, note that, to be consistent with the other plots in this year's report, the "box and whisker" charts on each plot show quartiles (last year's report showed 20th and 80th percentiles).

These program-level data are for unit costs that (unlike PAUC and APUC) are adjusted for any changes in procurement quantity. These results compare recurring procurement unit costs at the initially estimated quantities, extrapolating data if quantities have been reduced. This approach provides a superior way of comparing what the units would have cost if we had not changed quantities by, essentially, measuring the shift in the cost-versus-quantity procurement cost curve from planned to actual.[14] In other words, we measure changes in procurement cost at the currently planned quantity to be purchased (often lower than the initial) *and* assume that

---

[14]This basic approach for quantity adjustment is one of the standard techniques employed by the cost analysis community—see, for example, the discussions in Hough (1992), Arena et al. (2006, pp. 5–6), and Younossi et al. (2007, pp. 13-14).

the original planned quantity still was being purchased. This approach allows us to examine on a unit basis the cost of the capability to acquire those units regardless of whether we increased or decreased quantity. Of course, quantity decreases may be due to unit-cost increases, and this approach will show such cost increases clearly.

Similar to the prior RDT&E results, cost growths are highly skewed, with arithmetic means higher than the medians. As noted elsewhere for the contract-level data, the overall magnitudes of the cost growths are not nearly as large as those for RDT&E. Also, there is considerable variability in the production cost growth across the MDAP portfolio.

Figure 2-20 shows that quantity-adjusted unit-cost growth has shown no statistically significant changes over original MS B baselines. Figure 2-21 condenses the y-axis scale to show all outliers, and the table at the bottom identifies the five largest cost-growth programs for each year.

Figure 2-20. Program Procurement Cumulative Unit-Cost Growth (Quantity Adjusted) Over Original MS B Baseline: Central Quartiles (2001–2013)


Median	1%	6%	8%	7%	3%	4%	6%	9%	5%	8%	8%	6%	4%	6%
Mean	3%	24%	21%	18%	18%	26%	30%	29%	49%	24%	27%	28%	26%	28%

(differences are *not* statistically significant)
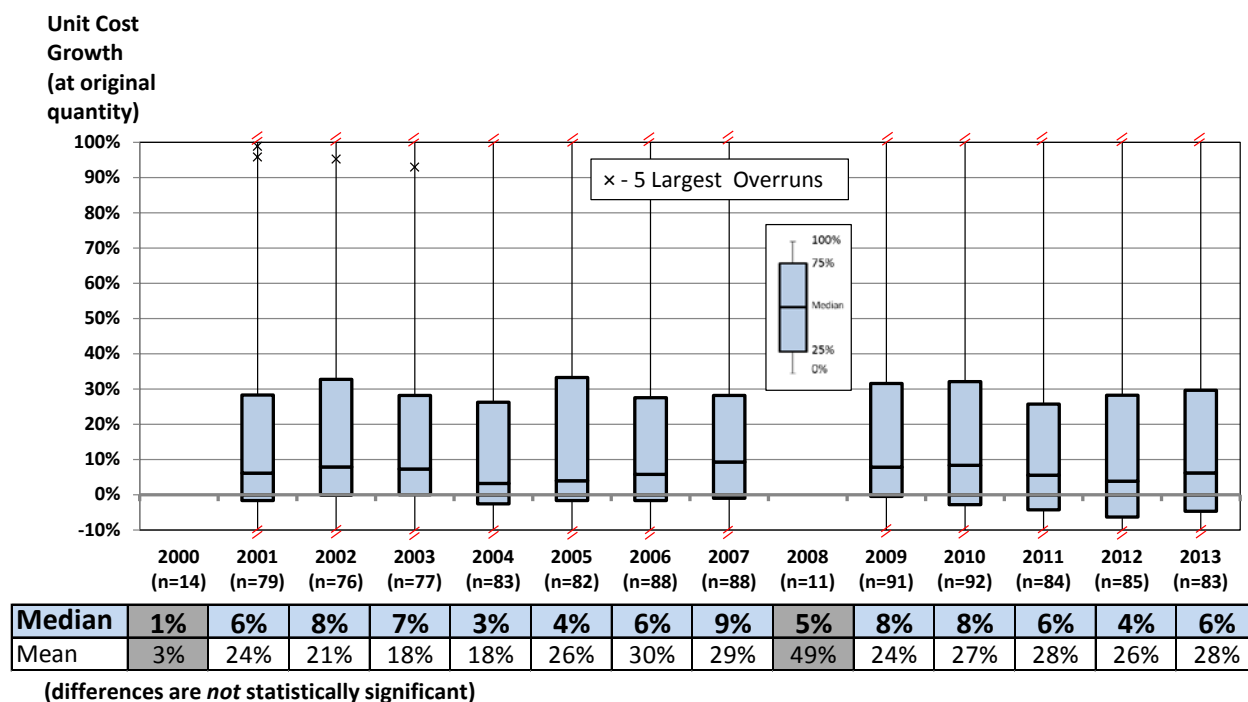
NOTE: Source budgetary cost data were reported in "base-year" dollars (adjusted for inflation). Boxes show second quartile, median, and third quartile; the min and max for each year are off the chart.

Figure 2-21. Program Procurement Cumulative Unit-Cost Growth (Quantity Adjusted) Over Original MS B Baseline: Outliers (2001–2013)
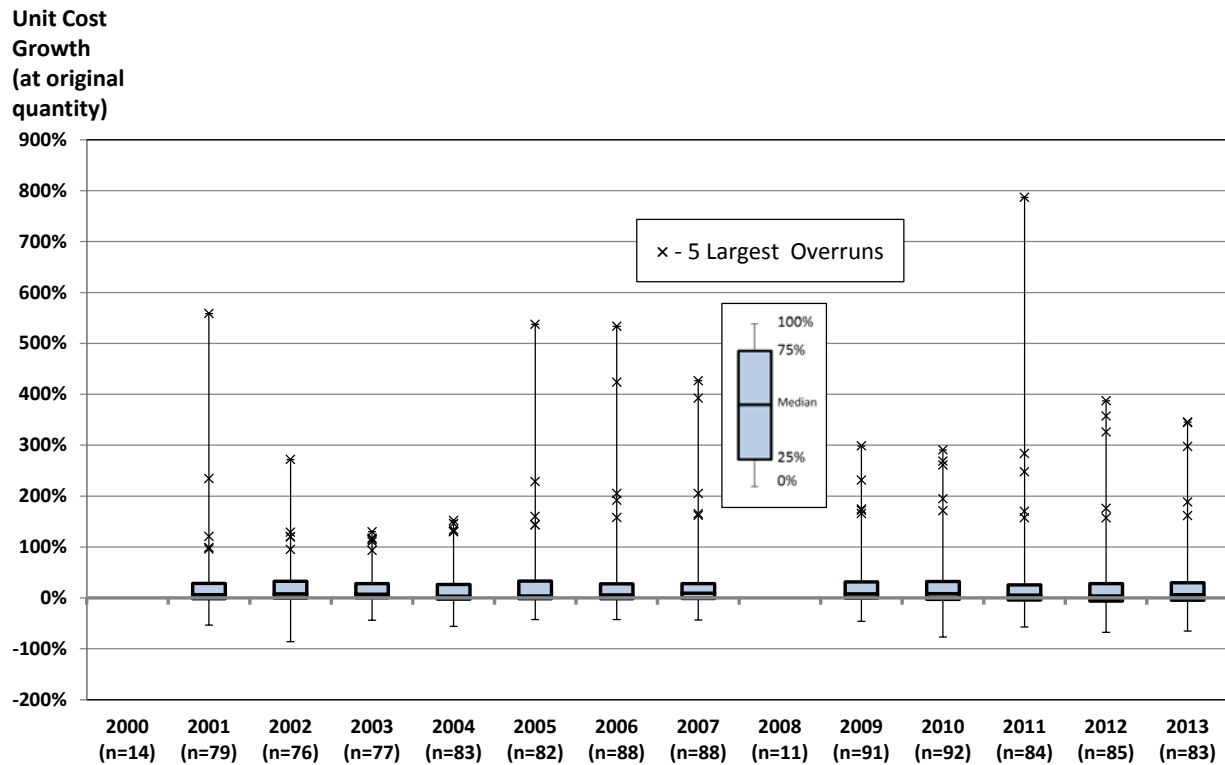
	2000 (n=14)	2001 (n=79)	2002 (n=76)	2003 (n=77)	2004 (n=83)	2005 (n=82)	2006 (n=88)	2007 (n=88)	2008 (n=11)	2009 (n=91)	2010 (n=92)	2011 (n=84)	2012 (n=85)	2013 (n=83)
Largest		SADARM Projectile	ATACMS BLK II/IIA	PAC-3 Fire Unit	SBIRS High			NPOESS		SBIRS High	ATIRCM QRC	C-130 AMP	EELV	
2nd Largest		ATIRCM CMWS	PAC-3 Fire Unit	GMLRS AW	EELV	NPOESS		SBIRS High		AEHF SV 1-4	SBIRS High		AEHF SV 1-4	
3rd Largest		PAC-3 Fire Unit	ATACMS BAT P3I	CH-47F		GMLRS AW	EELV			H-1 Upgrades	AEHF SV 1-4		SBIRS High	
4th Largest		GMLRS AW		EELV	GMLRS AW	EELV	GMLRS AW			C-130 AMP		GMLRS AW		
5th Largest		CH-47F	H-1 Upgrades	SBIRS High	PAC-3 Fire Unit	CH-47F		C-130 AMP		GMLRS AW	EFV	CH-47F		H-1 Upgrades

NOTE: Source budgetary cost data were reported in "base-year" dollars (adjusted for inflation). Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Figure 2-22 shows quantity-adjusted cost growth in overlapping 2-year periods. The last three periods show recent improvements, although the last period was only lower than two prior periods. The earliest period (1999–2001) was statistically higher than all (unshaded) periods. Figure 2-23 condenses the y-axis scale to show all outliers, and the table at the bottom identifies the five largest cost-growth programs for each year.

Figure 2-22. Program Procurement Cumulative Unit-Cost Growth (Quantity Adjusted) Over 2-Year Spans: Central Quartiles (1999–2013)

Unit Cost
Growth
(at orginal quantity)
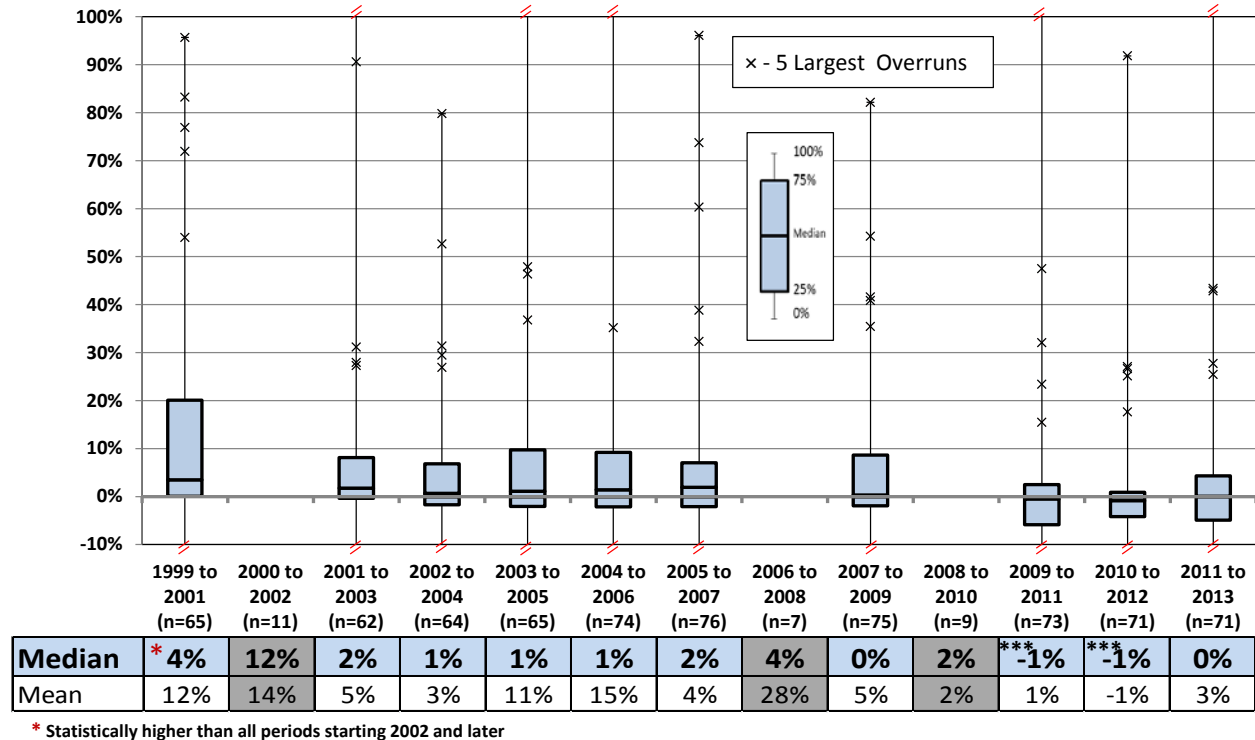
Median	*4%	12%	2%	1%	1%	1%	2%	4%	0%	2%	***-1%	***-1%	0%
Mean	12%	14%	5%	3%	11%	15%	4%	28%	5%	2%	1%	-1%	3%

Columns: 1999 to 2001 (n=65), 2000 to 2002 (n=11), 2001 to 2003 (n=62), 2002 to 2004 (n=64), 2003 to 2005 (n=65), 2004 to 2006 (n=74), 2005 to 2007 (n=76), 2006 to 2008 (n=7), 2007 to 2009 (n=75), 2008 to 2010 (n=9), 2009 to 2011 (n=73), 2010 to 2012 (n=71), 2011 to 2013 (n=71)

* Statistically higher than all periods starting 2002 and later

NOTE: Source budgetary cost data were reported in "base-year" dollars (adjusted for inflation). Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Figure 2-23. Procurement Program Cumulative Unit-Cost Growth (Quantity Adjusted) Over 2-Year Spans: Outliers (1999–2013)

Unit Cost
Growth
(at orginal quantity)
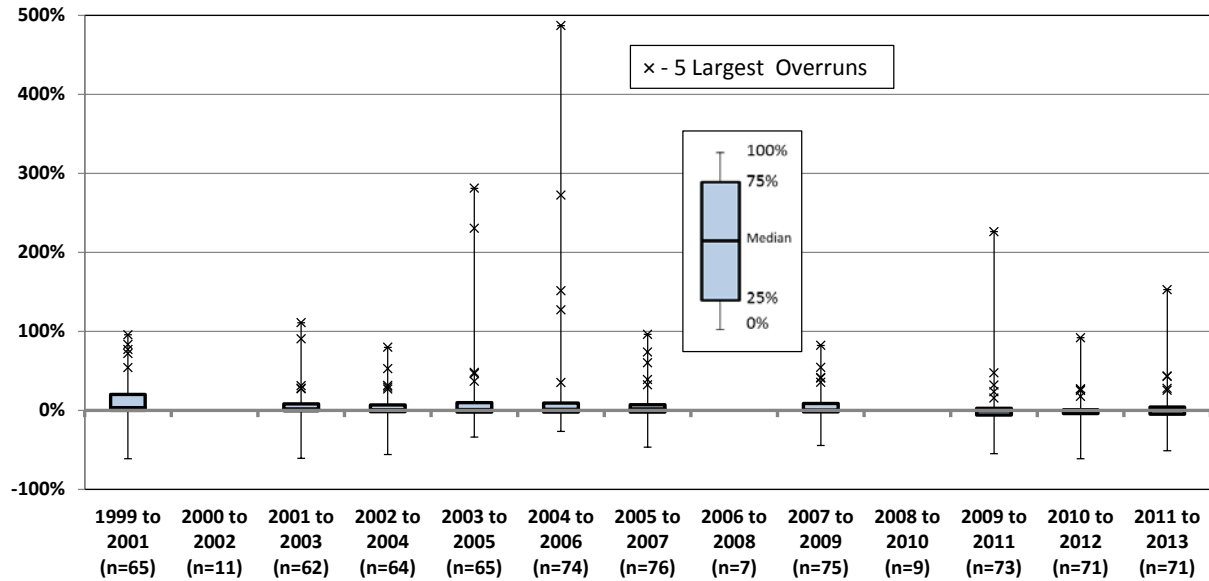
	Largest													
Largest	CH-47F		SBIRS High	EELV	NPOESS		C-130 AMP		B-2 RMP		C-130 AMP	JTRS HMS	JPALS Inc 1A
2nd Largest	ATIRCM CMWS		EELV	CH-47F	SBIRS High	LAND WARRIOR	EFV		AEHF SV 1-4		JTRS HMS	B-2 EHF Inc 1	VTUAV
3rd Largest	GMLRS AW		AEHF SV 1-4	SBIRS High	UH-60M	SBIRS High	NPOESS		VTUAV		Excalibur	AEHF SV 1-4	AH-64E Reman
4th Largest	MH-60R		TACTOM	AEHF SV 1-4	FCS	C-130 AMP	AEHF SV 1-4		AH-64E Reman		FAB-T	MQ-9 Reaper	AEHF SV 1-4
5th Largest	H-1 Upgrades		B-1B CMUP Computer	UH-60M	LAND WARRIOR	EFV	LHA 6		H-1 Upgrades		WGS	CH-53K	JTRS HMS

NOTE: Source budgetary cost data were reported in "base-year" dollars (adjusted for inflation). Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

COMPARISON OF RECENT DEVELOPMENT COST GROWTH TO PRIOR STUDIES

Given the recent cost-growth performance discussed above, we examined how they compare to historical performance as measured in prior studies.

Direct comparisons with or between any of these studies are difficult given wide variations in sampling and methodology. For example, Arena et al. (2006, pp. 7–12) summarizes reported cost growth for MDAP development and production in various studies. These ranged from SAR-based to contract-based analysis over different time periods. Most results were adjusted for inflation and quantity—but not all. Hough (1992), Jarvaise, Drezner, and Norton (1996), and Arena et al. (2006) note that differences in sample (including military departments, weapon system types, years, baseline points used, competed, sole sourced, and inclusion of incomplete programs) and analytic methodology (e.g., adjusting for inflation; adjusting for quantity; weighting by size) can have a dramatic effect on the results. Also, most reported means (averages) rather than medians: medians provide a better measure of central tendency when the distribution is skewed.

Nevertheless, McNicol (2005, p. 8) and Arena et al. (2006, p. 8) noted that total cost growth was much higher in the years prior to major acquisition initiatives such as those by Packard in 1969. For example, an unpublished RAND study of 24 weapon systems acquired in 1946–1959 found average total cost growth (development and production) of about 223 percent when adjusting for inflation and quantity.

In comparison, Table 2-9 summarizes the development and production cost growth measured by various studies and compares them to results from current SARs. Studies of completed and ongoing acquisitions as early as 1960 with widely varying samples reported average development cost growth ranging from 20 to 126 percent, average procurement cost growth ranging from 17 to 65 percent, and average total cost growth ranging from 14 to 110 percent. Note that some studies reported development average cost growth less than that for procurement, but most showed development cost growth to be higher than that for production.

In their own study, Arena et al. (2006) analyzed completed programs with SARs from 1968–2003 that were deemed "similar in type" to Air Force MDAPs (i.e., no ship or submarine MDAPs were included). They reported median development adjusted cost growth of 34 percent (average was 58 percent) and a median procurement adjusted cost growth of 40 percent (average was 44 percent).

In comparison, Figure 2-14 above shows the set of all current MDAPs reporting in SARs since 2001 reported median development cost growth of 4 to 19 percent (average is 48 to 87 percent) since their original MS B baseline (2001–2013 SARs). Although the current cost growths fit within the historical range from earlier studies, program completion and commodity bias makes this comparison tenuous.

When we looked only at the 88 completed MDAPs (i.e., those that stopped reporting in the SARs), the median total RDT&E cost growth dropped to 3 percent (mean was 76 percent).

Excluding ship programs yielded a median of 4 percent (mean of 80 percent) in development cost growth. This median was lower than that reported by Arena et al. (2006) for their subset of 46 completed programs from 1968–2003 that are "similar" to Air Force MDAPs, but the recent mean was higher. The dramatic difference between median and mean development cost growth on our recent MDAPs is the result of extreme outliers presented and discussed earlier with Figure 2-15. This shows again that the best measure of central tendency in skewed distributions is the median rather than the mean. This also indicates that relative to historical standards, MDAPs since 2001 have generally shown improvements but we continue to have unacceptably high variability with extreme outliers.

Table 2-9. Comparison of Recent MDAP Cost Growth to Historical Values

	Development		Production	
	median	*mean*	*median*	*mean*
Mixed MDAP sampling:				
Historical studies (1960–2003)		20–126 %		17–65 %
Recent (annually, 2001–2013)	4–19 %	48–87 %	3–9 %	18–30 %
Completed MDAPs (non-ship):				
Historical (MDAPs from 1968–2003) *	34 %	58 %	40 %	44 %
Recent (2001–2013) **	4 %	80 %	3 %	38 %

NOTES: *Arena et al., 2006. **n=82 for development and n=80 for production. Measures are adjusted for inflation. Production measures are adjusted for quantity changes. Current MDAP development growths are those since original MS B baseline (see Figure 2-14). Current MDAPs production growths are quantity-adjusted since original MS B baseline (see Figure 2-20). Most measures were cited as arithmetic means, which are sensitive to the values of extreme outliers but are not the best measure of central tendency on skewed distributions such as those generally found on acquisition programs. Historical measures are those cited in Arena et al., 2006, pp. 9–10.

Also, Figure 2-20 above shows the set of all current MDAPs reporting in SARs since 2001 reported median production cost growth of 3 to 9 percent (average is 18 to 30 percent) since their original MS B baseline (2001–2013 SARs). Again, while in the historical ranges of earlier studies, program completion and commodity bias makes this comparison very tenuous.

Similarly, when we just looked at the 87 completed MDAPs (i.e., those that stopped reporting in the SARs), the median quantity-adjusted unit procurement cost growth dropped to 2 percent (mean was 35 percent). Excluding ship programs yielded a median of 3 percent (mean of 38 percent) in procurement cost growth. Here, both the procurement median and mean were lower than that reported by Arena et al. (2006) for their subset of 44 completed programs from 1968–2003 that are "similar" to Air Force MDAPs.

In general, our recent results are encouraging relative to prior studies of historical program-cost performance. However, because the studies have widely varying sampling and analytic methodologies, direct comparisons to the larger set of historical studies remain imprecise.

CONTRACT-LEVEL TRENDS IN COST AND PRICE GROWTH: DEVELOPMENT

We examined recent trends in cost and price growth on development contracts. Figure 2-24 shows the real cost growth (adjusted for inflation) on 83 MDAP contracts in our dataset since the year 2000; Figure 2-25 shows the price-growth trend. Contracts are plotted by their start date, and a simple linear regression trend line is shown. Most of these contracts have been nearly completed, but some larger programs (e.g., KC-46A and F-35) were included with cost and price to date given their magnitude and to show how they compare to the broader performance.

The trend lines are slightly downward, but the variation is very high and significant outliers remain. Thus, as with last year's USD(AT&L) report, there are indications of improvement at the contract level, but the variation remains significant and the outlier cases must be prevented going forward. These figures also show how cost and price tend to be similar.

Figure 2-24. MDAP Contract Cost-Growth Trend by Start Date: Development (2000–2013)
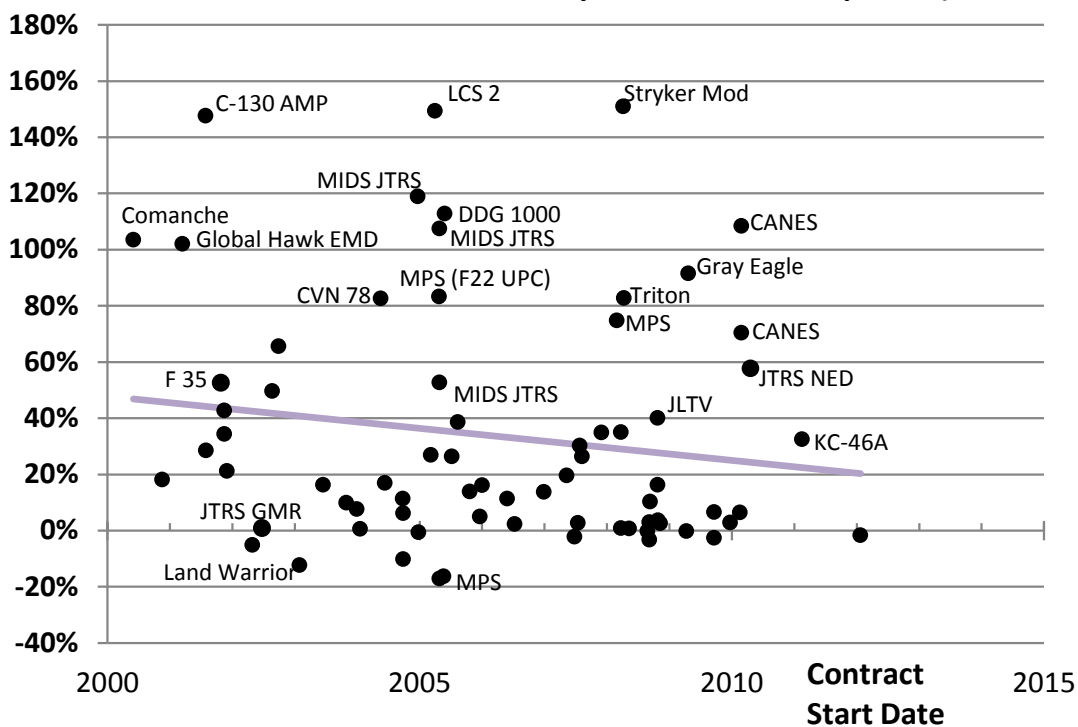
NOTE: Cost growth is in real terms (i.e., adjusted for inflation). Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Figure 2-25. MDAP Contract Price-Growth Trend by Start Date: Development (2000–2013)
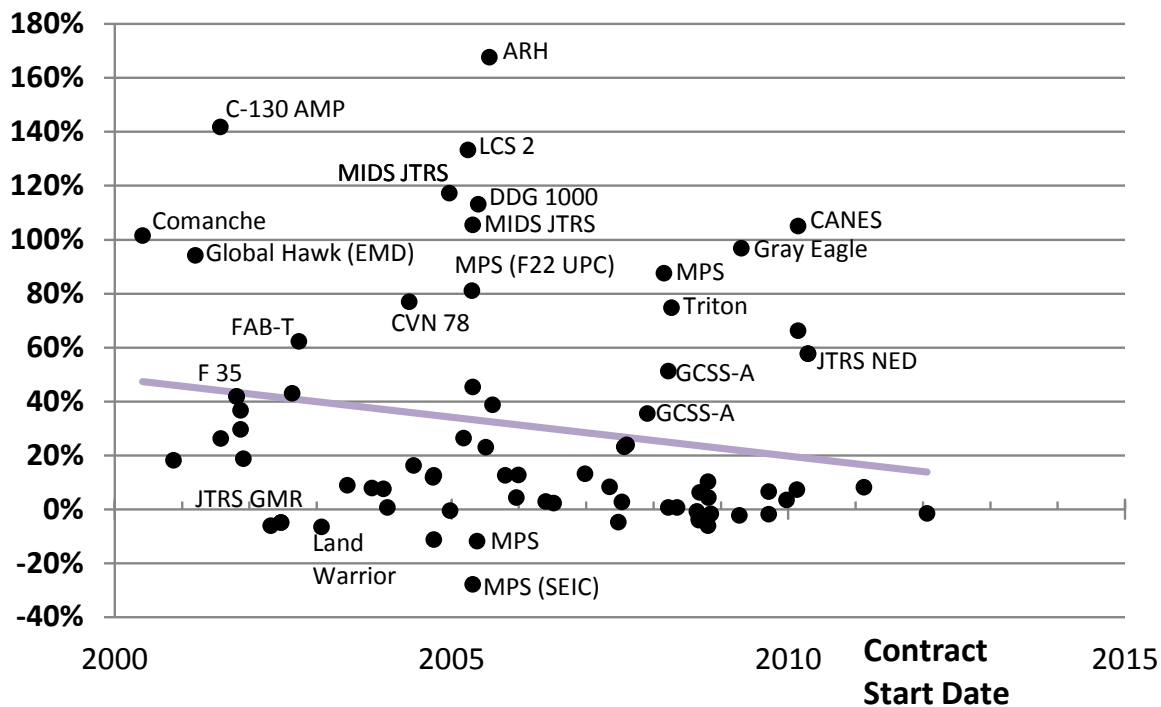
NOTE: Price growth is in real terms (i.e., adjusted for inflation). Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

CONTRACT-LEVEL TRENDS IN CONTRACT COST AND PRICE GROWTH: PRODUCTION

We examined recent trends in cost and price growth on production contracts. Figure 2-26 shows the real cost growth (i.e., adjusted for inflation) on 83 MDAP contracts in our dataset since 2003 (the few earlier contracts in our dataset were rejected by statistical tests—see Appendix A). Figure 2-29 shows the price-growth trend. Contracts are plotted by their start date, and a simple linear regression trend line is shown.

Unlike in development, price growth in real terms is flat. As expected from last year's analysis of cost growth, cost and price growth in production is generally lower than in development. Here it is just above zero when adjusting for inflation. (Recall that last year's report showed early production cost growth on contracts to be about 9 percent in *unadjusted* then-year dollars, see USD[AT&L] report, 2013). We do find some large outliers in production, however. These cases will be investigated to determine root causes for the large cost growth.

Figure 2-26. MDAP Contract Cost-Growth Trend: Production (2003–2013)
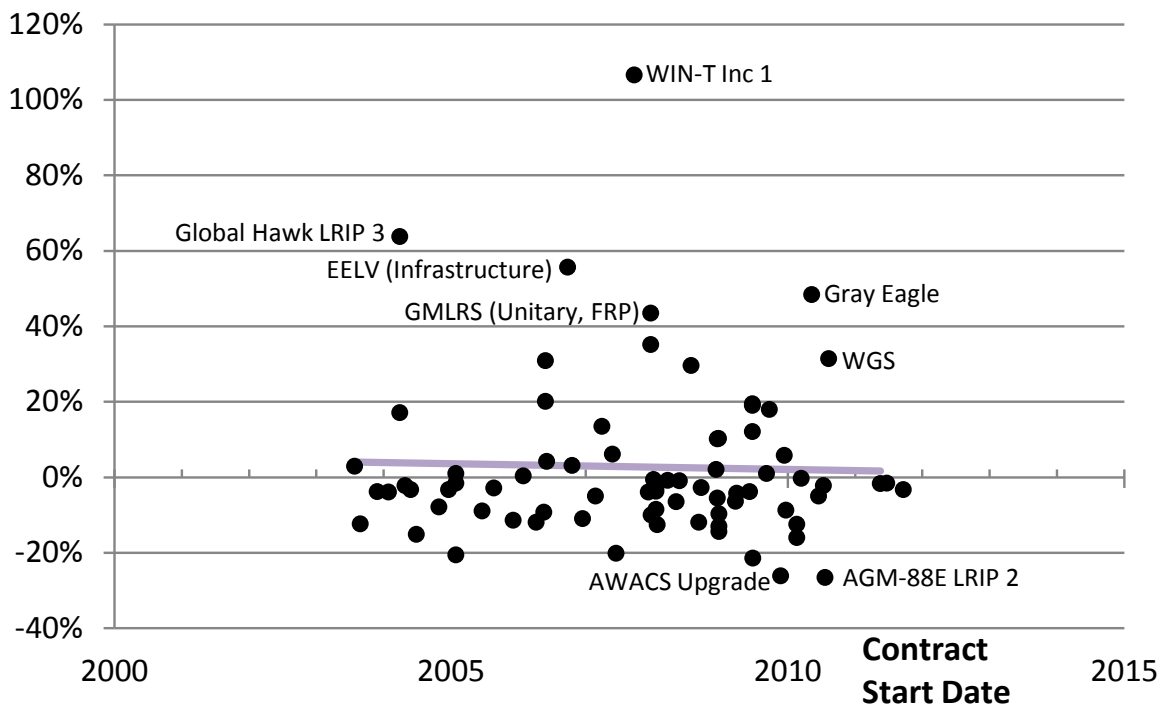
NOTE: Cost growth is in real terms (i.e., adjusted for inflation). Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Figure 2-27. MDAP Contract Price-Growth Trend: Production (2003–2013)
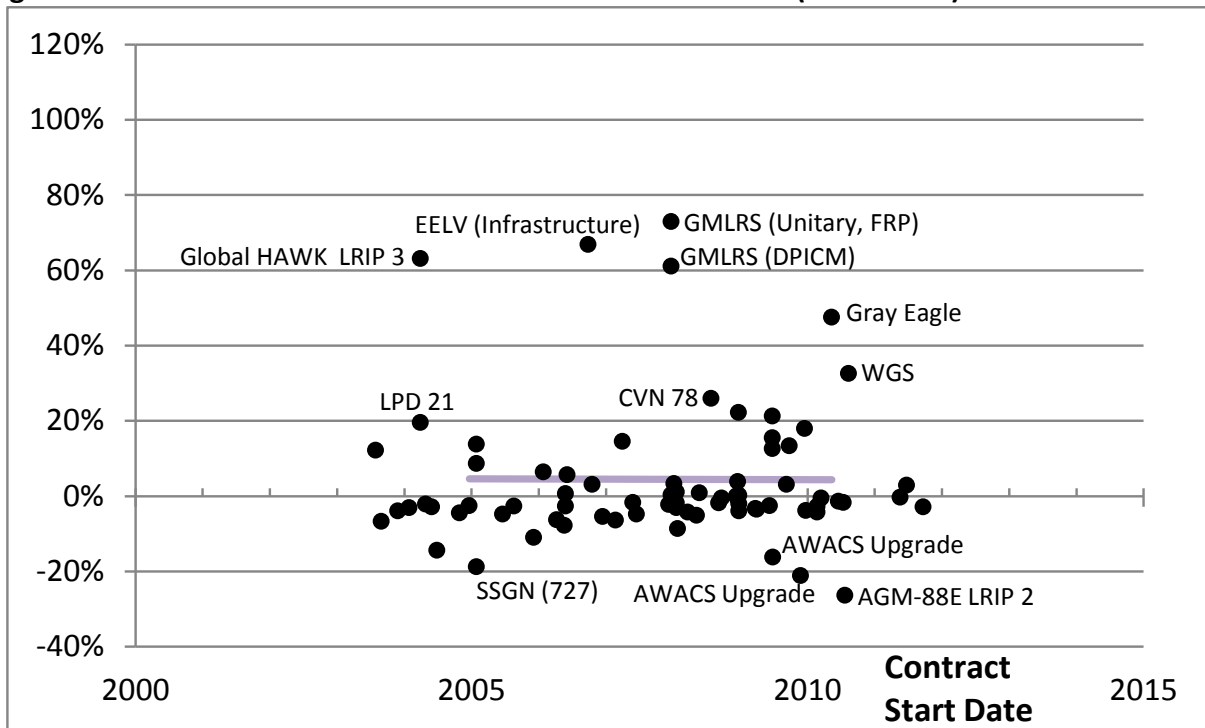
NOTE: Price growth is in real terms (i.e., adjusted for inflation). Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

INSTITUTIONAL ANALYSES: MILITARY DEPARTMENTS

This section summarizes various acquisition performance metrics analyzed by Component. Here we examined significant program-level cost growth as exhibited by Nunn-McCurdy breaches as well as contract-level cost, price, and schedule growth.

Nunn-McCurdy Cost Breaches

One measure of acquisition program cost performance is the Nunn-McCurdy breach rate by Component. Figure 2-28 shows Nunn-McCurdy breach counts by year from 1997 to 2013.

Figure 2-28. Nunn-McCurdy Breaches by Component per SAR Year (1997–2013)
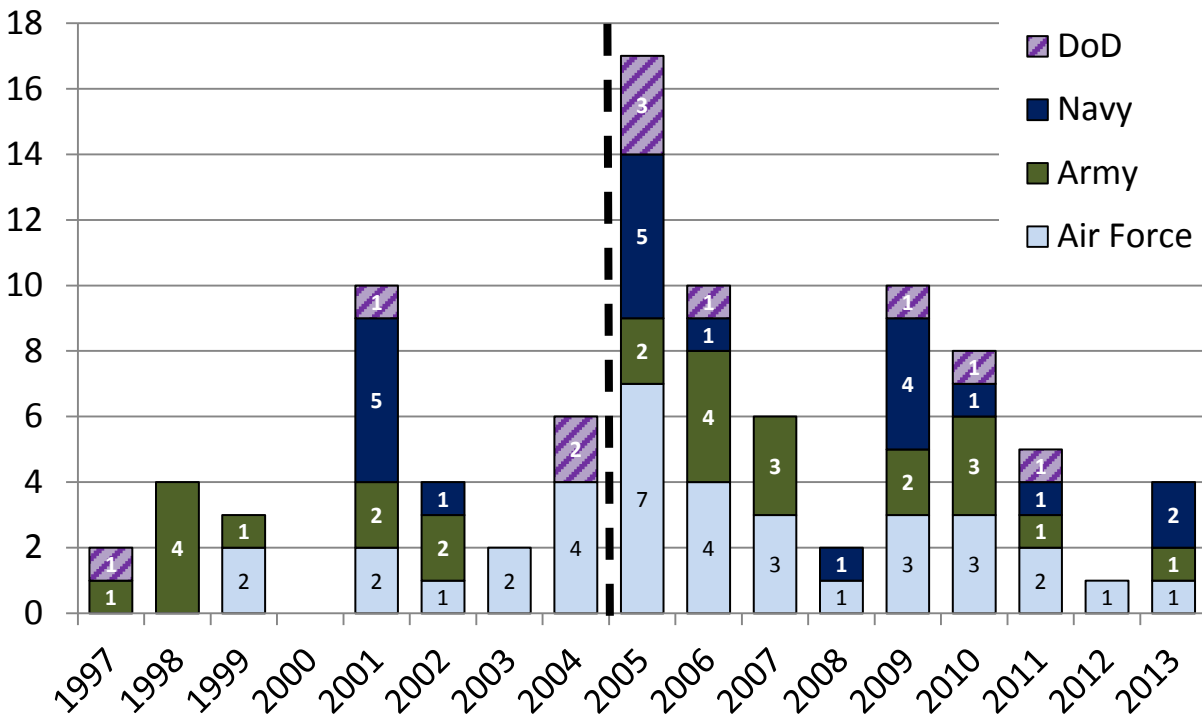
NOTE: The criteria for breaches were changed in NDAA 2006, affecting counts starting with 2005. Breaches are determined using "base-year" dollars (i.e., adjusting for inflation). This plot includes the number of breaches in each annual SAR reporting cycle, which nominally equates to calendar year but may include updates early in the following calendar year from the President's Budget Request. Breaches in different years for different thresholds or baselines for the same program are included in each respective year. If a program reported both a significant and critical breach in the same year, only one breach is shown here. Nunn-McCurdy breach reporting was established in the NDAA for FY1982, so the count shown here for 1997 may differ from that by others depending on whether prior notification for the same breach without rebaselining has occurred.

Table 2-10 below summarizes a different analysis of Nunn-McCurdy breaches by Component. Similar to the commodity summary above, we do not "double count" programs that have breached multiple times. This allows us to get a sense of the tendency of programs within each Component to breach.

Historically, about a third of MDAPs had at least one breach (i.e., about two-thirds have cost growth below 15 percent). Among the three military departments, the Air Force has the most critical breaches (total number and as a percentage), Army has the highest Component breach rate, and the Navy has the lowest breach rate. DoD programs have the highest breach rate at 58 percent; this is higher still than the 50 percent breach rate for programs labeled as Joint (DoD) in last year's USD(AT&L) report.[15] DoD programs also had the highest percentage of

---

[15] Note that there is a small sample set of only 12 DoD programs in this dataset.

programs that had a critical breach. At least two-thirds of programs that breach go critical (i.e., fewer remain at the significant level), except for Army programs, which are split.

All breaches are listed regardless of cause. If a program had both a significant and a critical breach, it was only included in the "programs with critical breach" column.

There are various causes of these breaches. Some programs may breach because of cancellation (e.g., Land Warrior) and some programs may have been canceled before their breach (e.g., VH-71). Last year's report provides a summary of the root-cause analyses to date. The analyses of the two critical breaches this year are ongoing and will be summarized in next year's report.

Table 2-10. Nunn-McCurdy MDAP Breach Rates by Component (1997–2013)

Component	Total # of Programs	# of Programs that Ever Breached	Breach Rate	# of Programs with at Most a Significant Breach	# of Programs with a Critical Breach
DoD	12	7	58%	2	5
Army	53	19	36%	9	10
Air Force	53	16	30%	4	12
Navy	63	16	25%	5	11
Total	**181**	**58**	**32%**	**20**	**38**

NOTE: The categorization of MDAPs by Component has been revised slightly since last year's USD(AT&L) report. If a program had both a significant and critical breach, it was only included in the "programs with critical breach" column. Not adjusted for quantity or other variances. "DoD" programs are programs categorized as DoD programs in SAR reporting, which include Joint programs and programs (such as Chemical Demilitarization) managed by an organization other than the Air Force, Army, or Navy. Breaches are determined using "base-year" dollars (i.e., adjusted for inflation).

Mission Effectiveness and Suitability of Acquired Systems: Military Departments

The following figures show the operational effectiveness and suitability of the systems (predominantly MDAPs) acquired by the Army, Navy, and Air Force as assessed by the DoD Director of Operational Test and Evaluation (DOT&E). Recall that these are binary ratings. The "Other DoD" category includes systems that are either joint and cannot be allocated to a single military department or are controlled by OSD or another entity. Given the low frequency of these programs, we grouped the ratings in three bins: programs before 2001, programs between 2001 and 2008, and programs since 2009. We chose these bins to establish a historical baseline, show performance during the years after the terrorist attacks of September 11, 2001, and have as recent a bin as possible that has sufficient numbers to avoid excessive sensitivity to sample count. Other groupings are possible.

Figure 2-29 shows the percentage of programs in each period rated as "effective." The most dramatic trend has been a decreasing trend for Other DoD programs, followed by a sizable drop

for the Navy. The Air Force, however, has shown recent improvements. Some caution is warranted about the precision of the percentages given the small sample sizes (especially for the Other DoD programs). Still, only three of the recent seven Other DoD programs were rated as effective at the time compared to all 2001–2008 programs being effective.

Figure 2-30 shows the percentage of programs in each period rated as "suitable." As expected given the low DoD-wide ratings shown earlier in Figure 2-3, these ratings are generally much lower than those for effectiveness. All three Components have remained about the same since 2009 as in the earlier prior, although the Air Force is much lower than the others. The Other DoD group has always been about as low as the Air Force is now. Again, some caution is warranted about the precision of the percentages given the small sample. For example, the seemingly slight improvement by the Army since 2009 is just due to one program rating of the 19 (i.e., each program contributes 5.2 percentage points, so if one more program was rated as not suitable in the last period the rating would fall below 70 percent).

Figure 2-29. Percent of Military Department Programs Rated Effective (1984–2013)
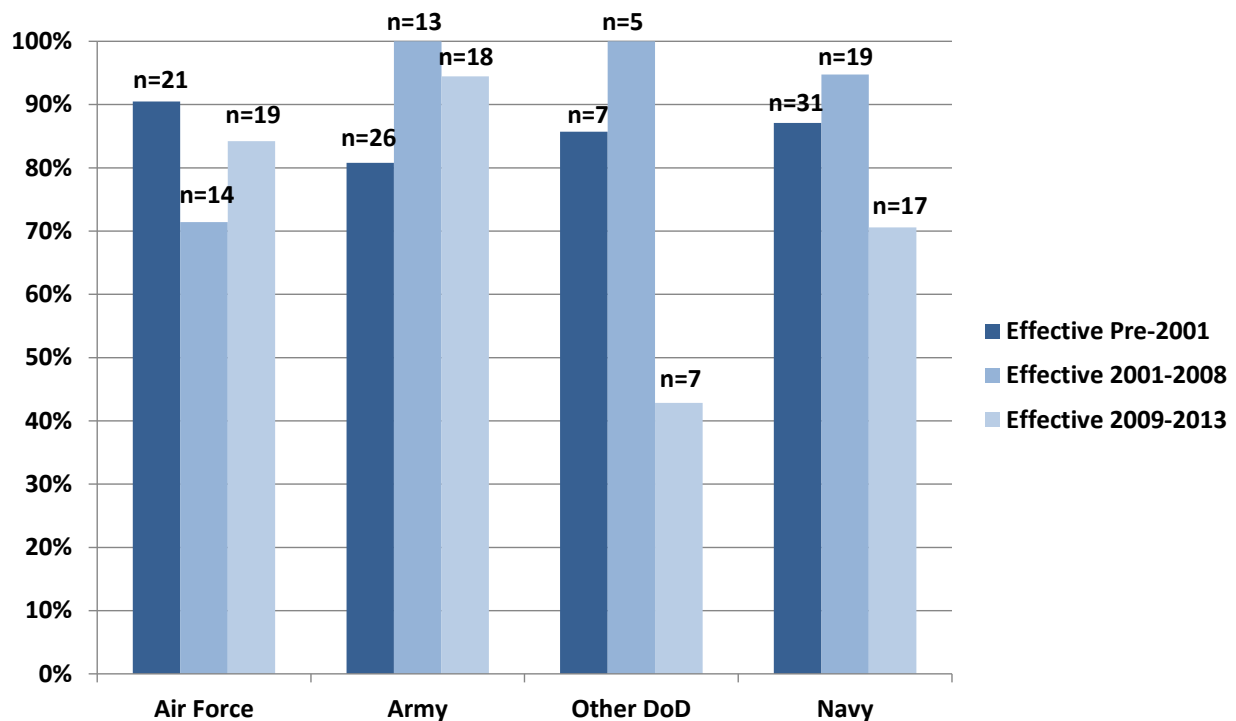
Source: DOT&E reports. NOTE: DoD programs were Joint or other programs that are not exclusive to a single Component. Sample sizes differ between effectiveness and suitability for some Components because there was not always a definitive binary judgment for effectiveness and suitability in all reports.

Figure 2-30. Percent of Military Department Programs Rated Suitable (1984–2013)
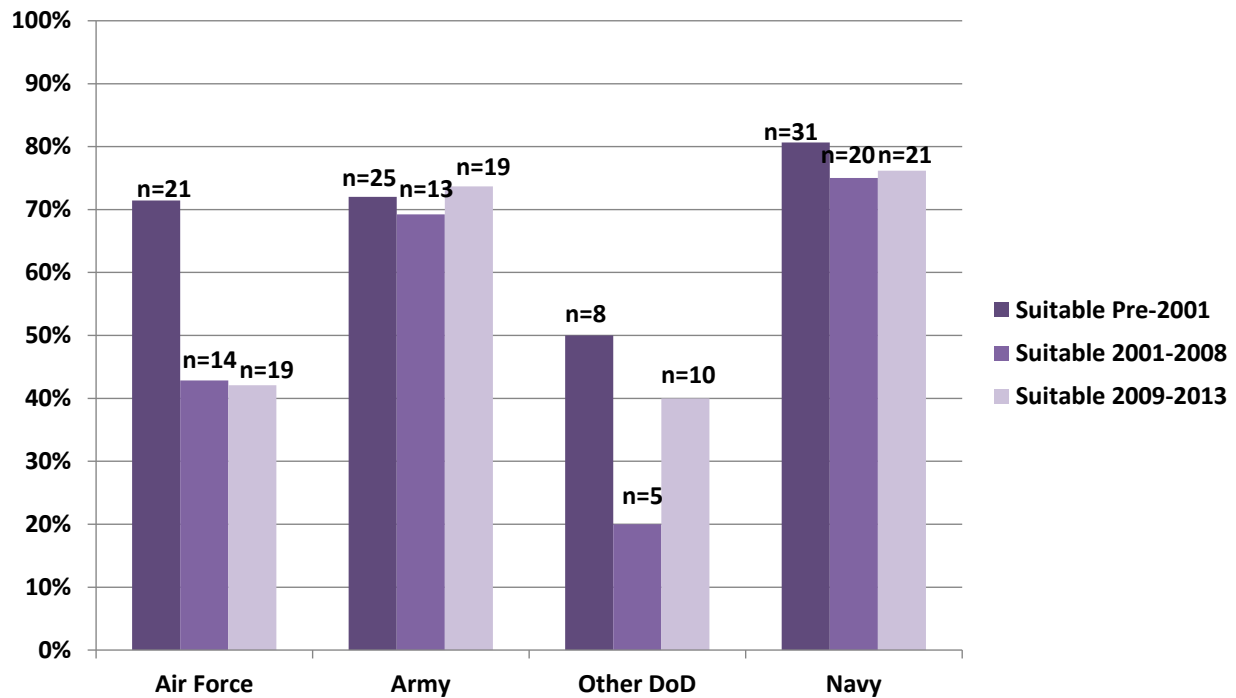
Source: DOT&E reports. NOTE: DoD programs were Joint or other programs that are not exclusive to a single Component. Sample sizes differ between effectiveness and suitability for some Components because there was not always a definitive binary judgment for effectiveness and suitability in all reports.

Contract Cost Growth and Service Acquisition Executives: Development

As with the earlier plots for the USD(AT&L), we updated the total cost growth on last year's set of development MDAP contracts in each Military Department shown by who was the Service Acquisition Executive (SAE) in office at each contract start date. A more extensive update of recent contracts will be included in next year's report. Some isolated patterns of interest emerged from prior analyses that are apparent in the following figures, but none are statistically significant unless they are explicitly as such.

Figure 2-31. Army Development Contract Total Cost Growth and SAE Tenures (1992–2014).
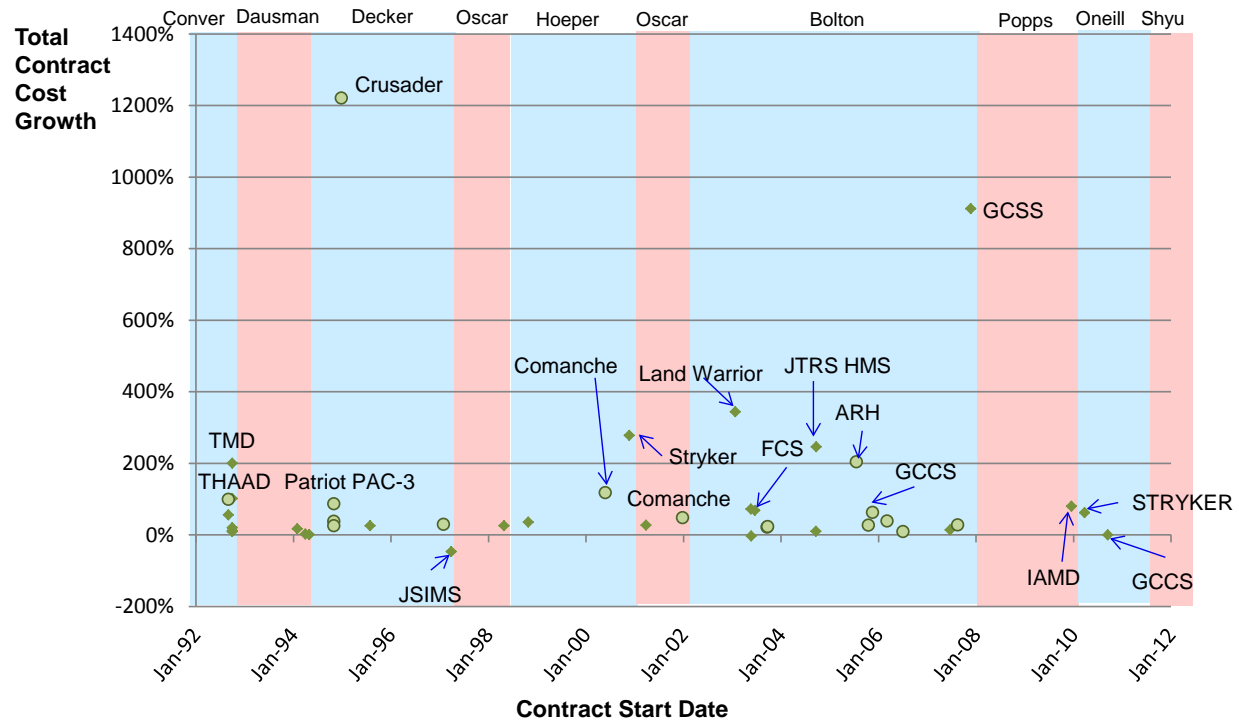
NOTE: Normally, contract start dates should be relatively close to prior major reviews (usually Milestone B decisions) by the SAE and MDA to approve contract award. Diamonds on the charts indicate growths mostly attributable to the start of the contract; circles represent significant work added later. Any white bars between SAE shaded regions represent periods where there was no confirmed executive. Source cost data were reported in "then-year" dollars (unadjusted for inflation).

Figure 2-32. Navy Development Contract Total Cost Growth and SAE Tenures (1992–2014).
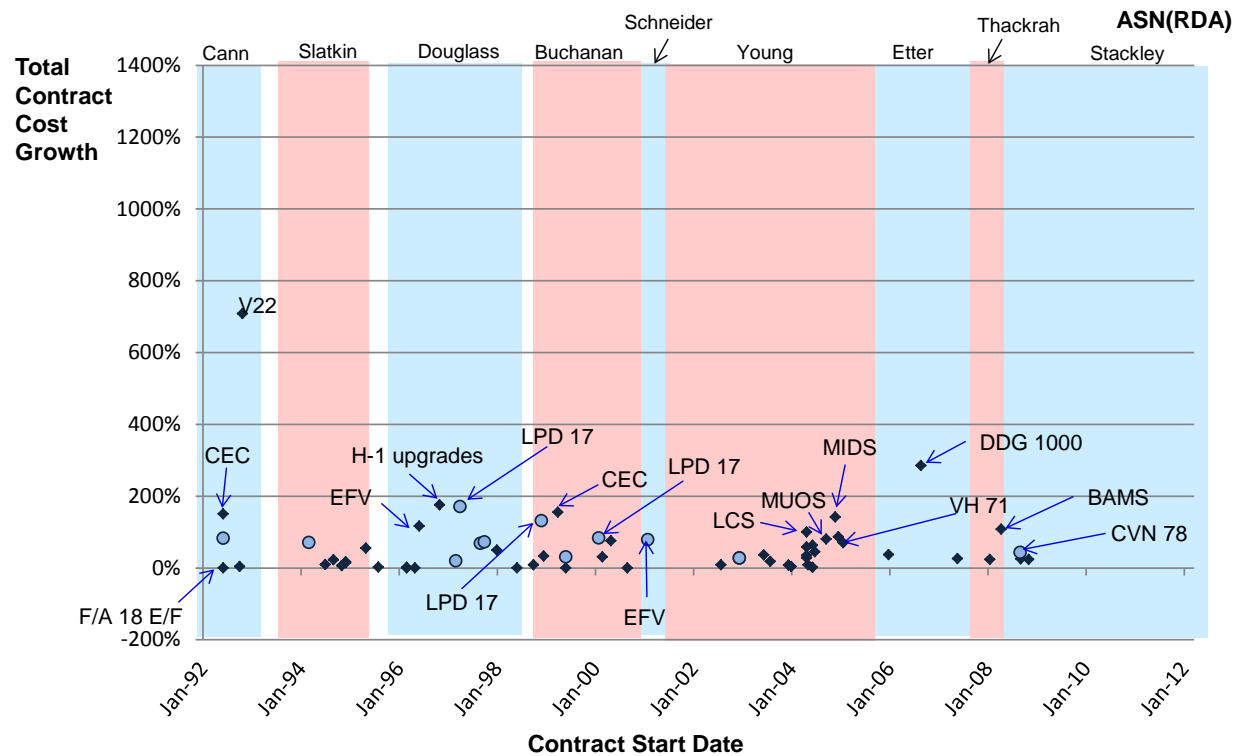
NOTE: Normally, contract start dates should be relatively close to prior major reviews (usually Milestone B decisions) by the SAE and MDA to approve contract award. Diamonds on the charts indicate growths mostly attributable to the start of the contract; circles represent significant work added later. Any white bars between SAE shaded regions represent periods where there was no confirmed executive. Source cost data were reported in "then-year" dollars (unadjusted for inflation).

Figure 2-33. Air Force Development Contract Total Cost Growth and SAE Tenures (1992–2014).
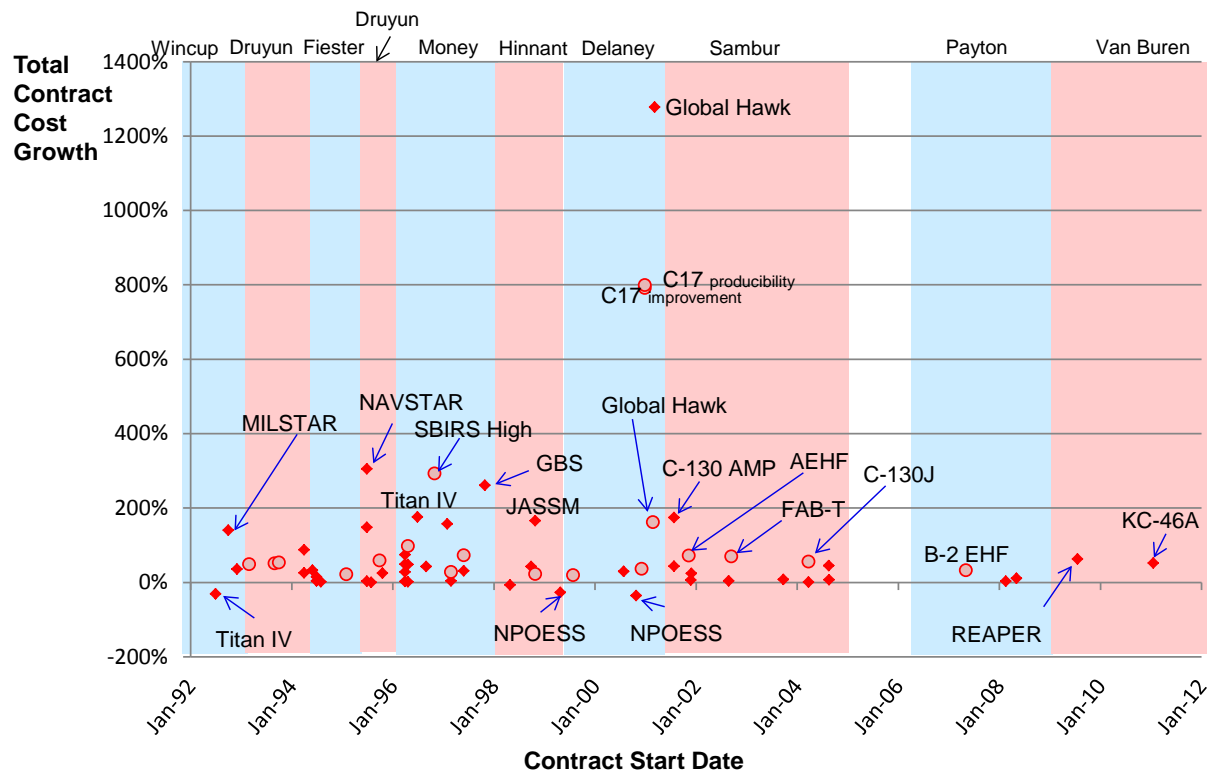
NOTE: Normally, contract start dates should be relatively close to prior major reviews (usually Milestone B decisions) by the SAE and MDA to approve contract award. Diamonds on the charts indicate growths mostly attributable to the start of the contract; circles represent significant work added later. Any white bars between SAE shaded regions represent periods where there was no confirmed executive. Source cost data were reported in "then-year" dollars (unadjusted for inflation).

These figures seem to imply visually some relationships between sitting SAE for major reviews before these contracts and the eventual performance of those contracts. For example, development contract schedule growth analysis shows that Navy growths appear to be better controlled since 2006 (Etter, Thackrah, and Stackley eras). However, such visual observations are subjective at this point rather than tested statistically, especially because these plots do not control for other variables that may be the dominate factors that affect cost growth on contracts.

However, while many visual observations of these figures are only apparent, some of our other statistical analyses are evident in these charts. For example, other analysis that found total cost growth for Air Force development contracts are lower by a statistically significant amount since 2002 can be seen visually in these charts (i.e., since the Sambur era).

Contract-Level Cost Growth by Military Department: Development

We examined contract-level cost growth in development by Military Department on prime contracts for MDAPs for which we have cost, price, and schedule growth data as well as final margins. Nearly all contracts shown are essentially complete (e.g., about 90 percent or more), but a small number of contracts such as the KC-46A were included, given their importance and the fact that sufficient time has elapsed on them to get early indications of their performance trend.

In addition to trend lines for cost growth, we also show trend lines for final margin on these contracts (see the margin scale on the right side of the plots) to see how final-margin trends compare to the cost-growth trends. Individual final margins are not shown. Trends are statistically significant only when indicated and are not weighted by spend. Standard tests were used to eliminate any outliers that unduly distort the underlying trends shown (see Appendix A). However, many outliers remain, and their position influences the apparent trend lines. Moreover, these plots show that the variation in cost and schedule growth is large; price-growth trends are not shown but are similar to that for cost growth.

In contract cost growth since year 2000, Army and Air Force trends appear downward while the Navy trend appears flat (none is statistically significant). Figure 2-34 shows that the Army's cost-growth trend is similar to that for all DoD contracts (including these Army contracts shown) but is shifted slightly lower. The trend for final margins on these contracts is essentially flat at about 6 percent.

Figure 2-35 shows that, in contrast to an apparently declining DoD-wide trend, the Department of the Navy's cost growth-trend is flat at about 40 percent. The final-margin trend for the Navy is apparently declining, which may reflect an increased effort to penalize poor cost performance. The F-35 was not included in the Navy set and is somewhat above the Navy trend line.

Figure 2-36 shows that the Air Force's cost growth-trend is apparently declining faster than the DoD-wide trend. This is apparently influenced by the two early large outliers in 2001. Like the Army, the Air Force final-margin trend is essentially flat. The F-35 was not included in the Air Force set, yet it lies on the Air Force trend line.

**Figure 2-34. Army Development Contract Cost-Growth and Final-Margin Trends (2000–2012)**
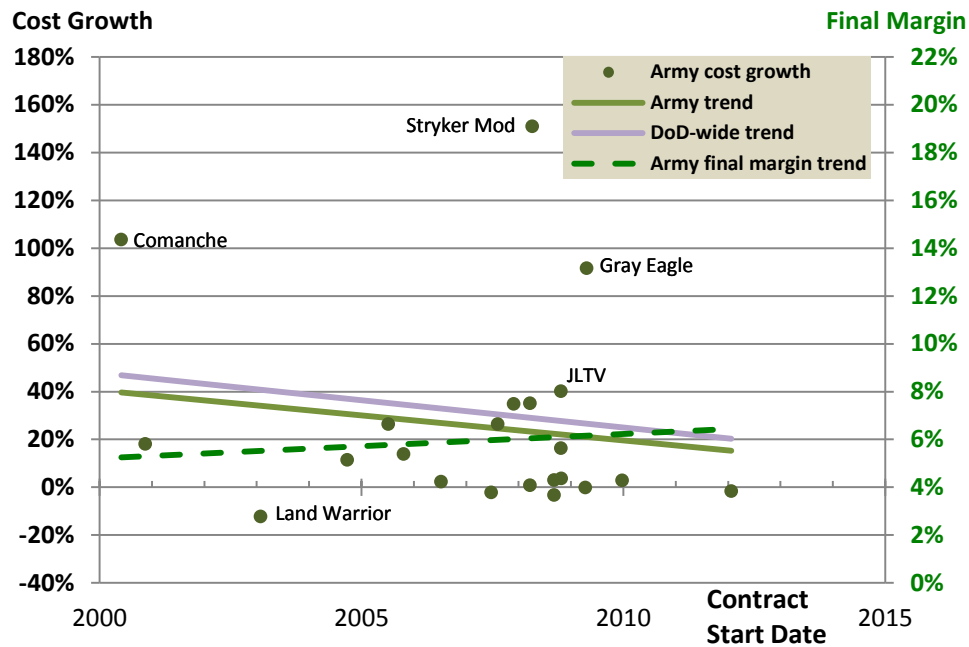
NOTE: Only the green dashed trend line is final margin, read from the right-hand scale. Individual data points and other trends are cost growth, read from the left-hand scale. The DoD-wide trend is for all contracts (including the Army contracts shown) and shown for comparison to the Army trend during the same time period. The Army trend is not steep enough to be statistically significant. Cost growth is in real terms (i.e., adjusted for inflation). Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Figure 2-35. Navy Development Contract Cost-Growth and Final-Margin Trends (2001–2010)
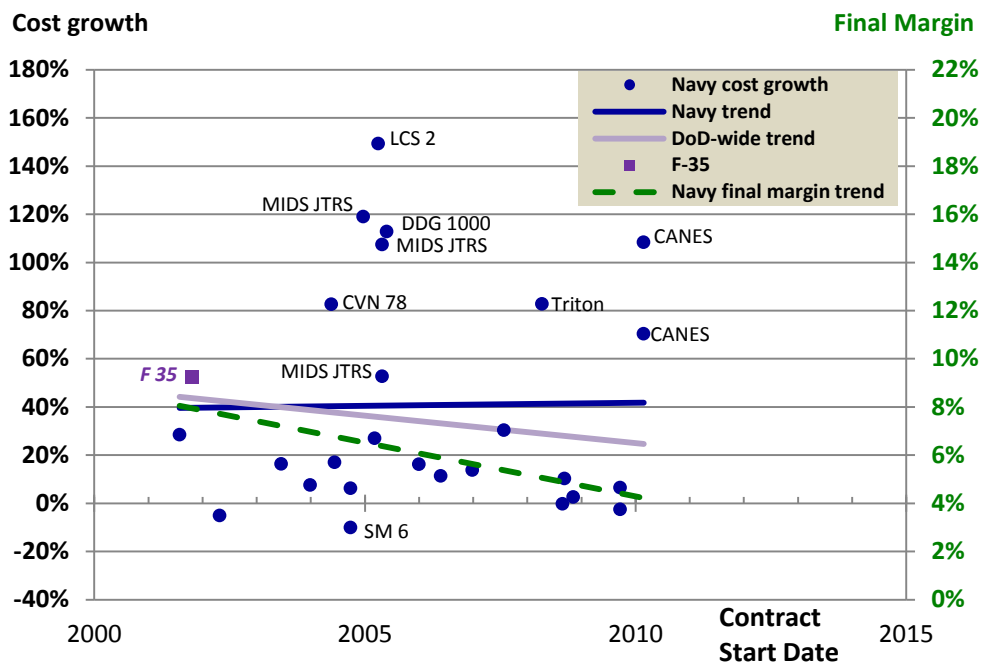
NOTE: Only the green dashed trend line is final margin, read from the right-hand scale. Individual data points and other trends are cost growth, read from the left-hand scale. The DoD trend is for all contracts (including the Navy contracts shown) and shown for comparison to the Navy trend during the same time period. The Navy trend is not steep enough to be statistically

significant. Cost growth is in real terms (i.e., adjusted for inflation). Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A). The F-35 JSF contract was not included in the Navy trend but is shown for comparison given its significance.

Figure 2-36. Air Force Development Contract Cost-Growth and Final-Margin Trends (2001–2011)
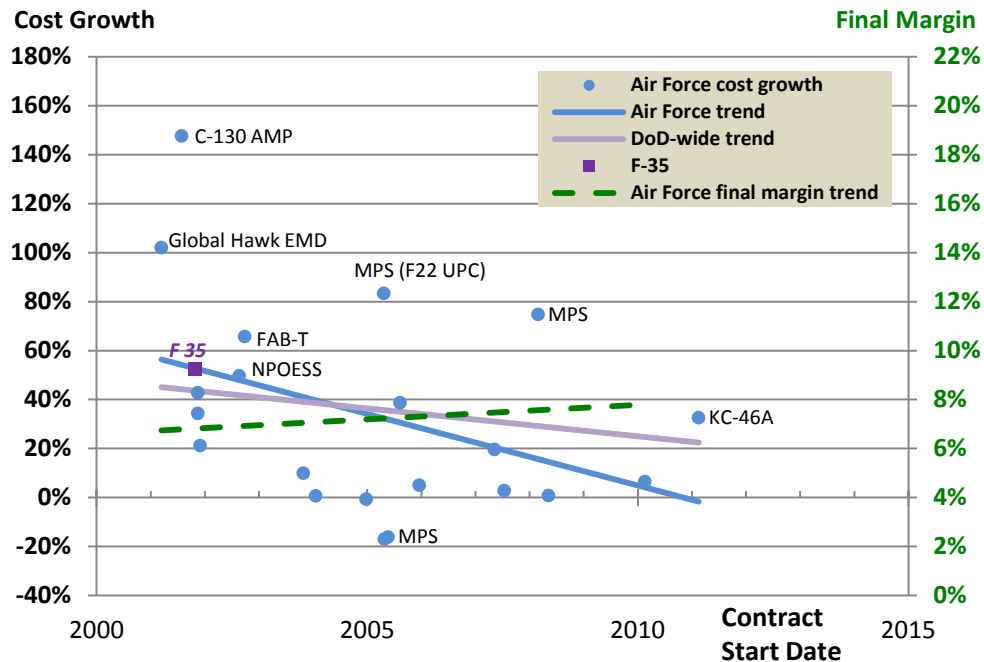
NOTE: Only the green dashed trend line is final margin, read from the right-hand scale. Individual data points and other trends are cost growth, read from the left-hand scale. The DoD trend is for all contracts (including the Air Force contracts shown) and shown for comparison to the Air Force trend during the same time period. The Air Force trend is not steep enough to be statistically significant. Cost growth is in real terms (i.e., adjusted for inflation). Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A). The F-35 JSF contract was not included in the Air Force trend but is shown for comparison given its significance.

Contract-Level Schedule Growth by Military Department: Development

As with cost growth, we examined schedule growth by Military Department on the same set of prime MDAP contracts discussed above. Standard tests were used to eliminate any outliers that unduly distort the underlying trends shown (see Appendix A). All three departments had downward trends in schedule growth since 2000, and all three of these trends are statistically significant. Many contracts had no schedule growth, but many outliers with positive schedule growth remain.

The Army trend starts higher than the DoD-wide trend but has a steeper slope (see Figure 2-37). The Navy trend tracked the DoD-wide downward trend (see Figure 2-38). As with development cost growth, the F-35 was not included in the Navy set and was slightly higher than the Navy trend. The significant Air Force trend had the same downward slope as the DoD-wide trend and is shifted about a half year shorter (see Figure 2-39). Again, F-35 was not included and was a year higher than the Air Force trend.

Figure 2-37. Army Development Contract Schedule-Growth Trends (2000–2012)
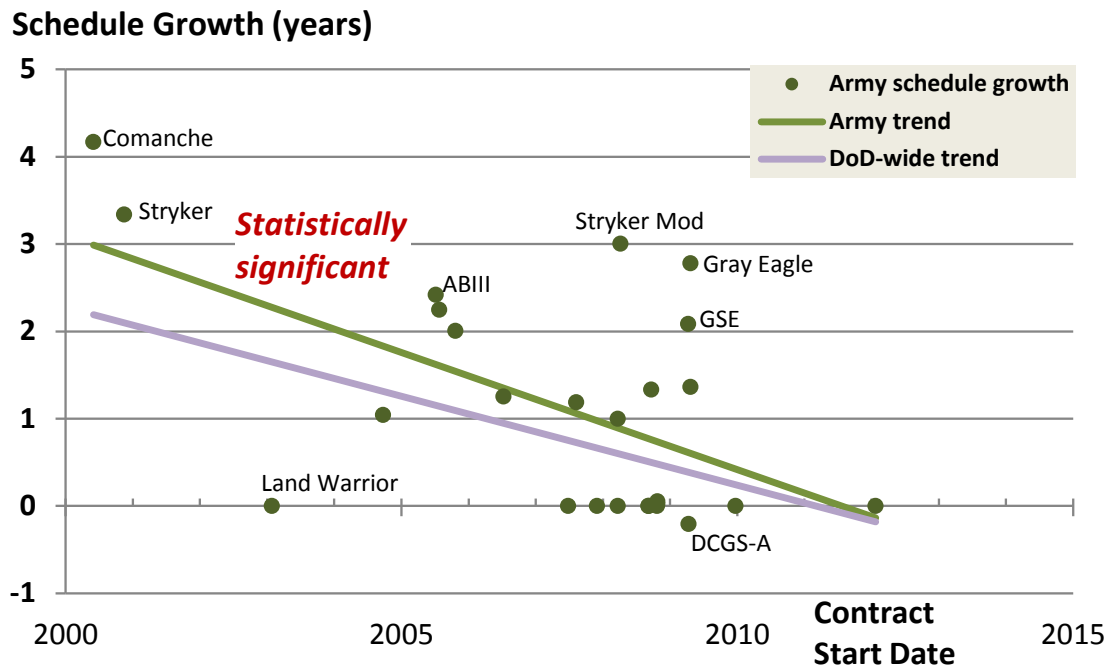
NOTE: The Army trend is statistically significant. The DoD trend is for all contracts (including the Army contracts shown) and shown for comparison to the Army trend during the same time period. Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Figure 2-38. Navy Development Contract Schedule-Growth Trends (2001–2010)
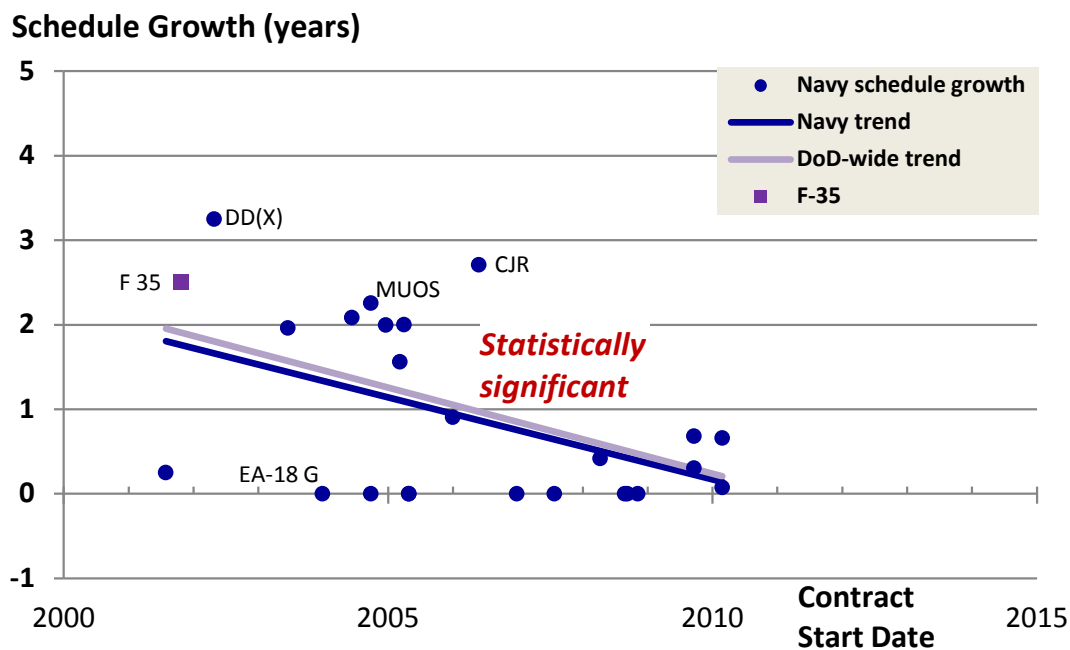
NOTE: The DoD trend is for all contracts (including the Navy contracts shown) and shown for comparison to the Navy trend during the same time period. The Navy trend is statistically significant. Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Figure 2-39. Air Force Development Contract Schedule-Growth Trends (2001–2011)
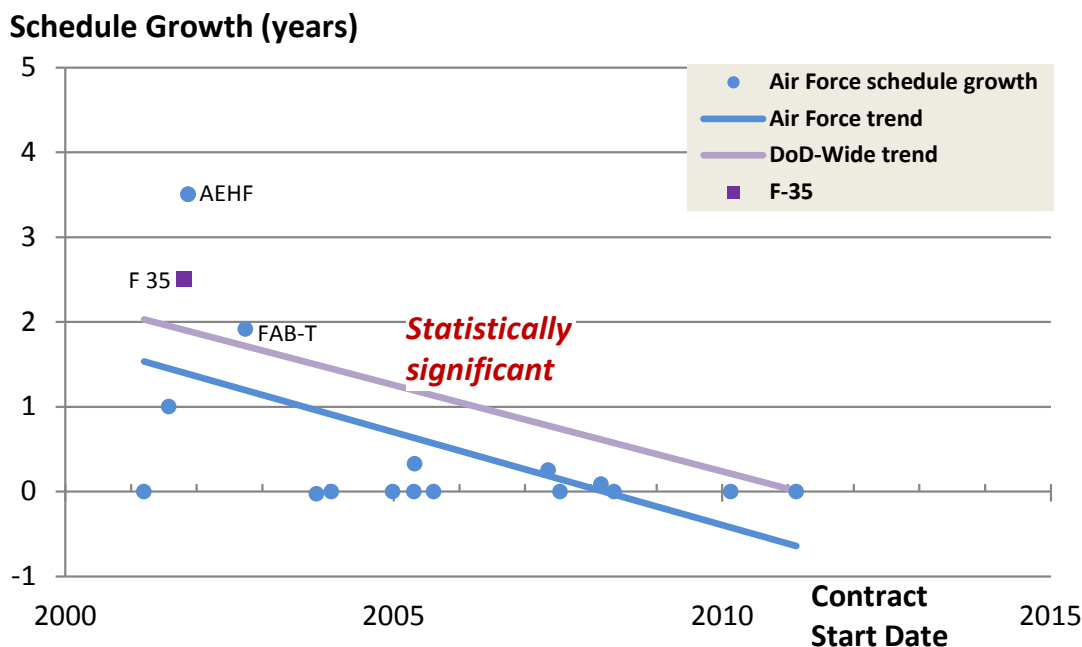
Schedule Growth (years)

NOTE: The DoD trend is for all contracts (including the Air Force contracts shown) and shown for comparison to the Air Force trend during the same time period. The Air Force trend is statistically significant. Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Contract Cost Growth and Service Acquisition Executives: Early Production

Below are updates of the total cost growth on last year's set of development MDAP contracts in each Military Department shown by who was the Service Acquisition Executive (SAE) in office at each contract start date. A more extensive update of recent contracts will be included in next year's report. Some isolated patterns of interest emerged from prior analyses that are apparent in the following charts. Unless shown, there was no noteworthy pattern observed.

Figure 2-40. Army Early Production Contract Total Cost Growth and SAE Tenures (1992–2014).
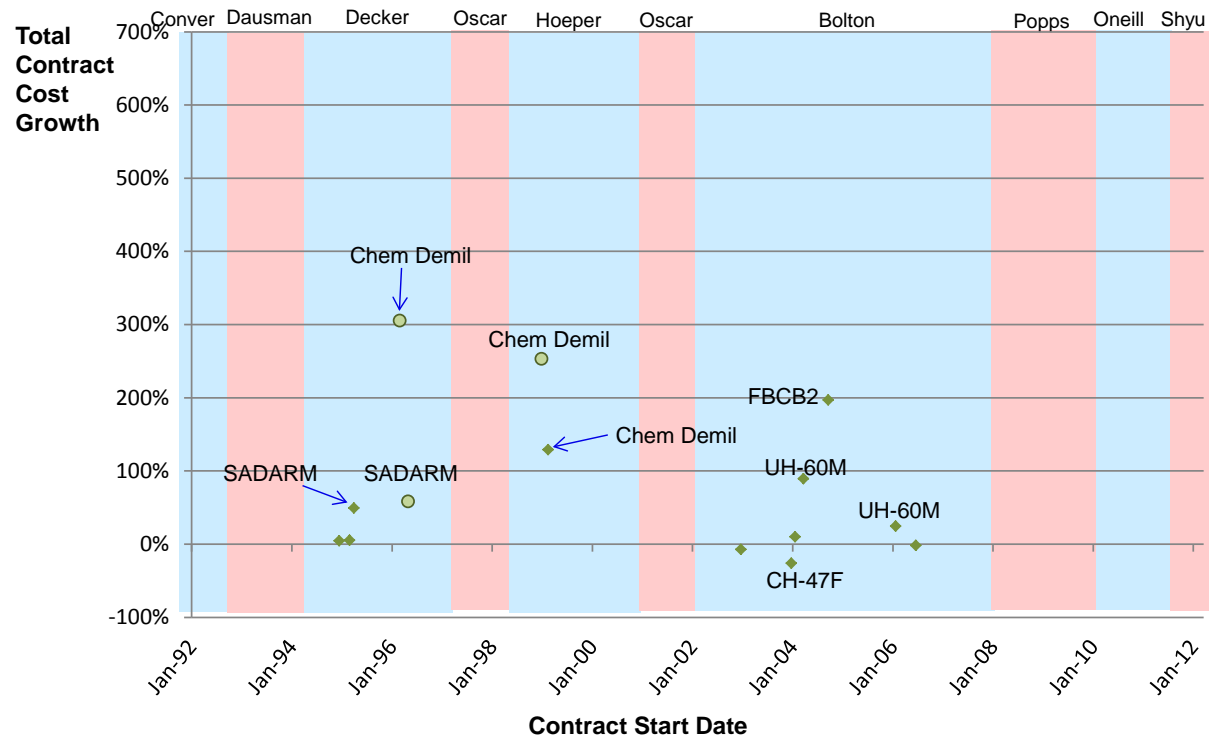
NOTE: Normally, contract start dates should be relatively close to prior major reviews (usually Milestone B decisions) by the SAE and MDA to approve contract award. Diamonds on the charts indicate growths mostly attributable to the start of the contract; circles represent significant work added later. Any white bars between SAE shaded regions represent periods where there was no confirmed executive. Source cost data were reported in "then-year" dollars (unadjusted for inflation).

Figure 2-41. Navy Early Production Contract Total Cost Growth and SAE Tenures (1992–2014).
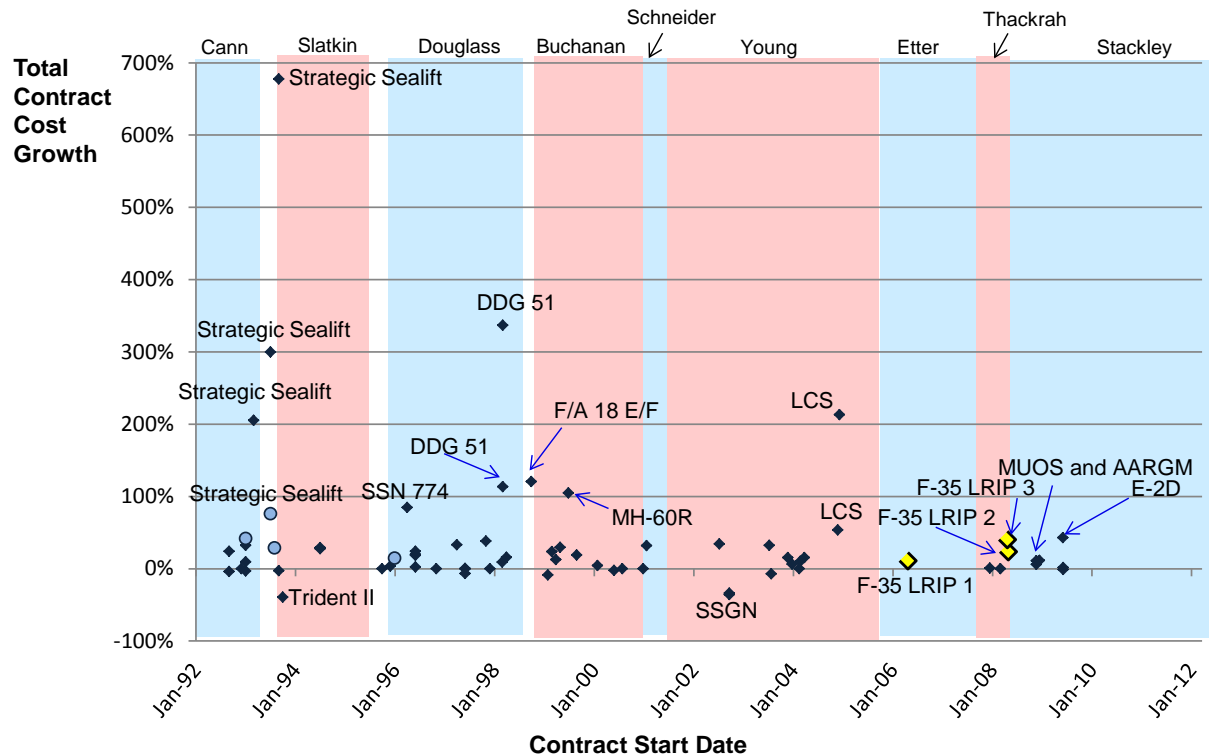
NOTE: Normally, contract start dates should be relatively close to prior major reviews (usually Milestone B decisions) by the SAE and MDA to approve contract award. Diamonds on the charts indicate growths mostly attributable to the start of the contract; circles represent significant work added later. Any white bars between SAE shaded regions represent periods where there was no confirmed executive. Source cost data were reported in "then-year" dollars (unadjusted for inflation).

Figure 2-42. Air Force Early Production Contract Total Cost Growth and SAE Tenures (1992–2014).
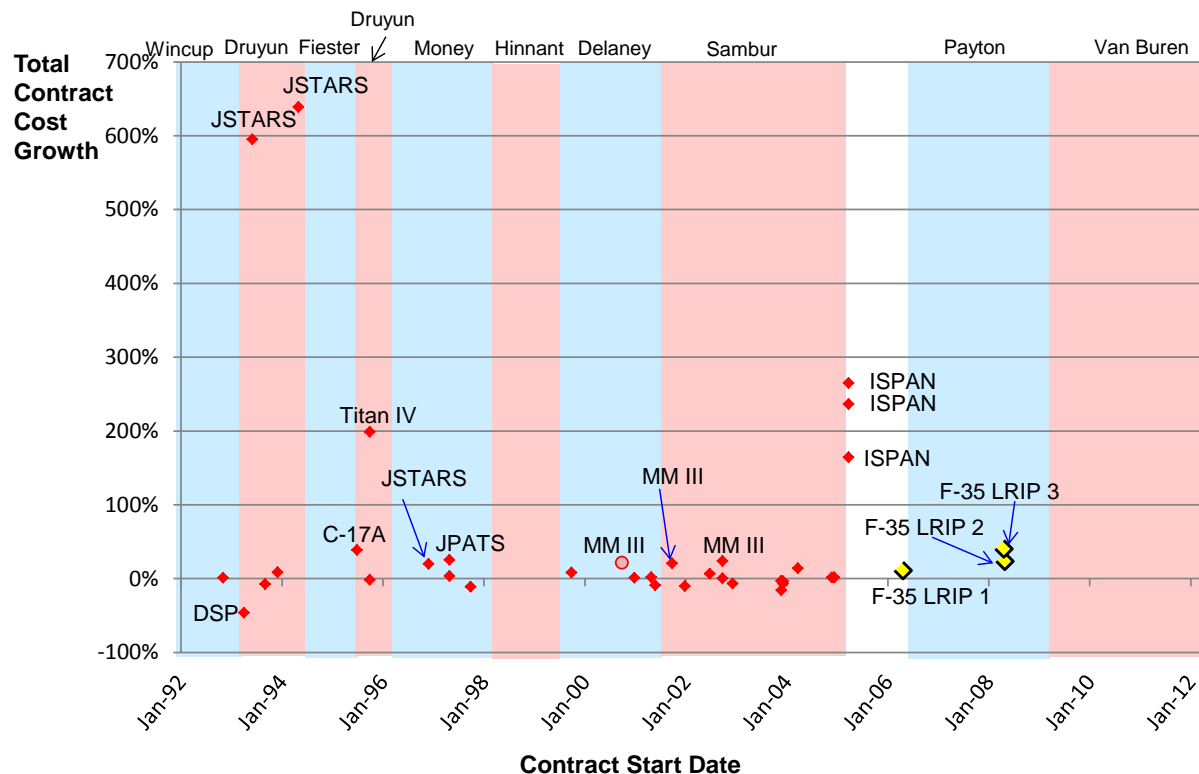
NOTE: Normally, contract start dates should be relatively close to prior major reviews (usually Milestone B decisions) by the SAE and MDA to approve contract award. Diamonds on the charts indicate growths mostly attributable to the start of the contract; circles represent significant work added later. Any white bars between SAE shaded regions represent periods where there was no confirmed executive. Source cost data were reported in "then-year" dollars (unadjusted for inflation).

These figures seem to imply visually some relationships between SAE in place for major reviews before these contracts and the eventual performance of those contracts. For example, since 2008 (Thackrah and Stackley eras), Navy production total cost growth appears better controlled than earlier Navy contracts. However, all such visual observations are subjective at this point rather than tested statistically, especially because these plots do not control for other variables that may be the dominate factors that affect cost growth on contracts.

Contract-Level Cost Growth by Military Department: Production

We now examine contract-level cost growth in production by Military Department on prime MDAP contracts for which we have cost, price, and schedule growth data as well as final margins. Standard tests were used to eliminate any outliers that unduly distort the underlying trends shown (see Appendix A). Since year 2000, the Army appears to have an upward trend while the Navy appears to be flat and the Air Force trending downward. None of these cost-growth trends was statistically significant.

Figure 2-43 shows the apparent Army cost-growth trend along with the apparent trend on final margin. Here final margins are much higher than those on development contracts. The final-margin trend is fairly flat at about 17 percent and is driven by the dominance of FFP contracts in the Army sample.

Figure 2-44 shows that the apparent Navy cost-growth trend in development is slightly downward and below zero. The apparent Navy final-margin trend is upward, which is not necessarily a concern, given that the cost performance is improving. Overall, the Navy's final margins are about 12 percent, which is much lower than the Army margins. This is due to the dominance of FPIF LRIP contracts in the Navy production dataset.

Finally, Figure 2-45 shows that the apparent Air Force cost growth is downward. Final margins are also slightly downward and are also lower than the Army and even Navy levels at about 10 percent.

Figure 2-43. Army Production Contract Cost-Growth and Final-Margin Trends (2004–2010)
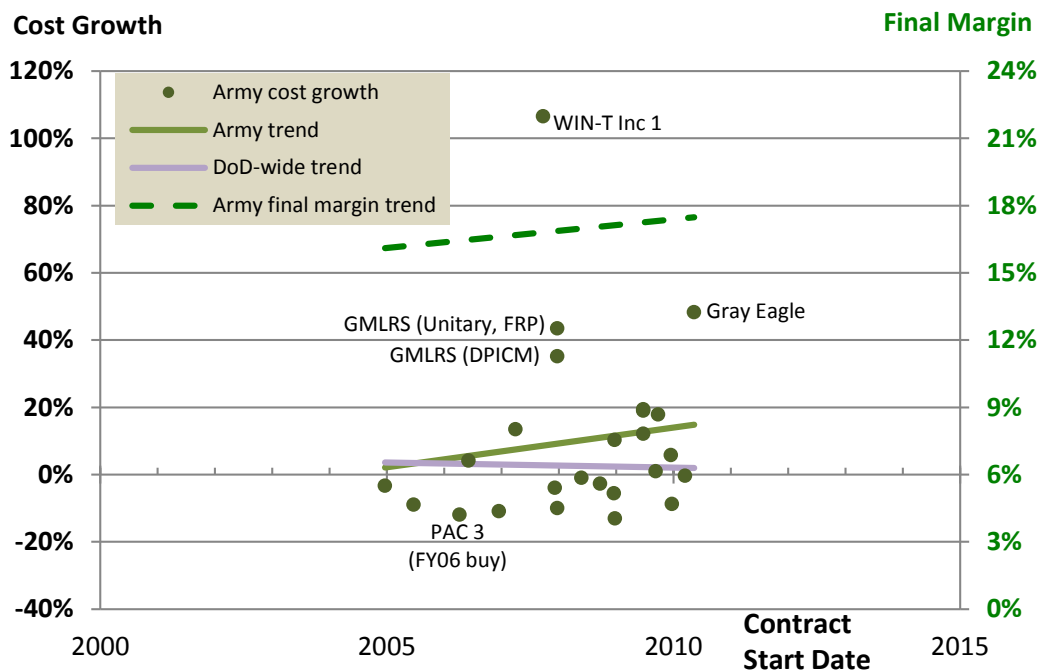
NOTE: Only the green dashed trend line is final margin, read from the right-hand scale. Individual data points and other trends are cost growth, read from the left-hand scal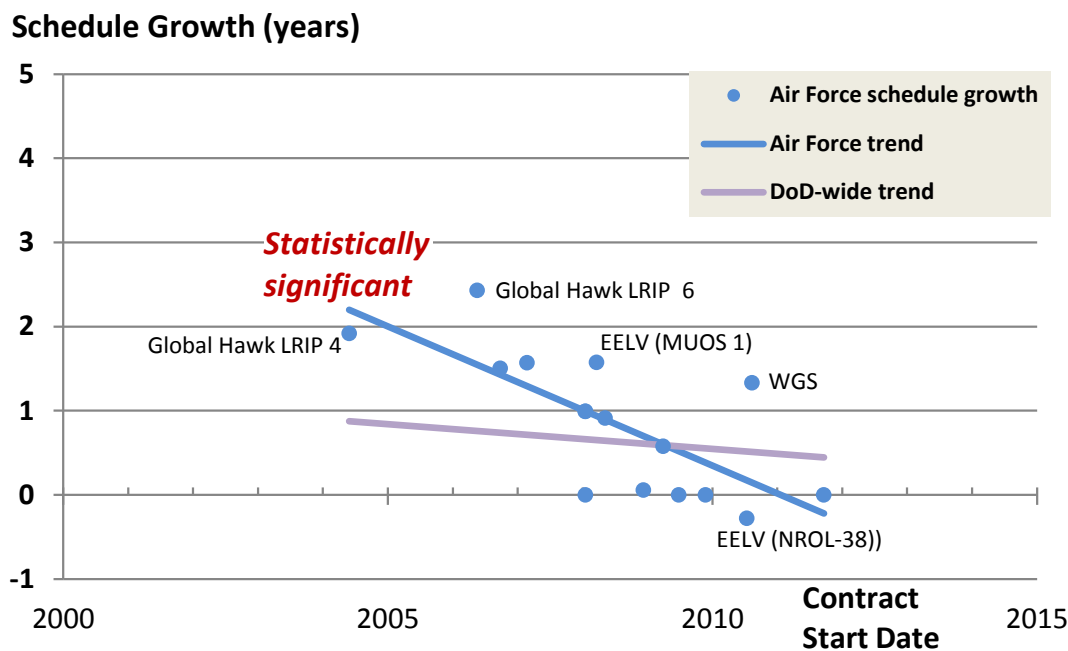e. The DoD trend is for all contracts and shown for comparison to the Army trend during the same time period. Cost growth is in real terms (i.e., adjusted for inflation). The Army trend is statistically significant. Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Figure 2-44. Navy Production Contract Cost-Growth and Final-Margin Trends (2003–2012)
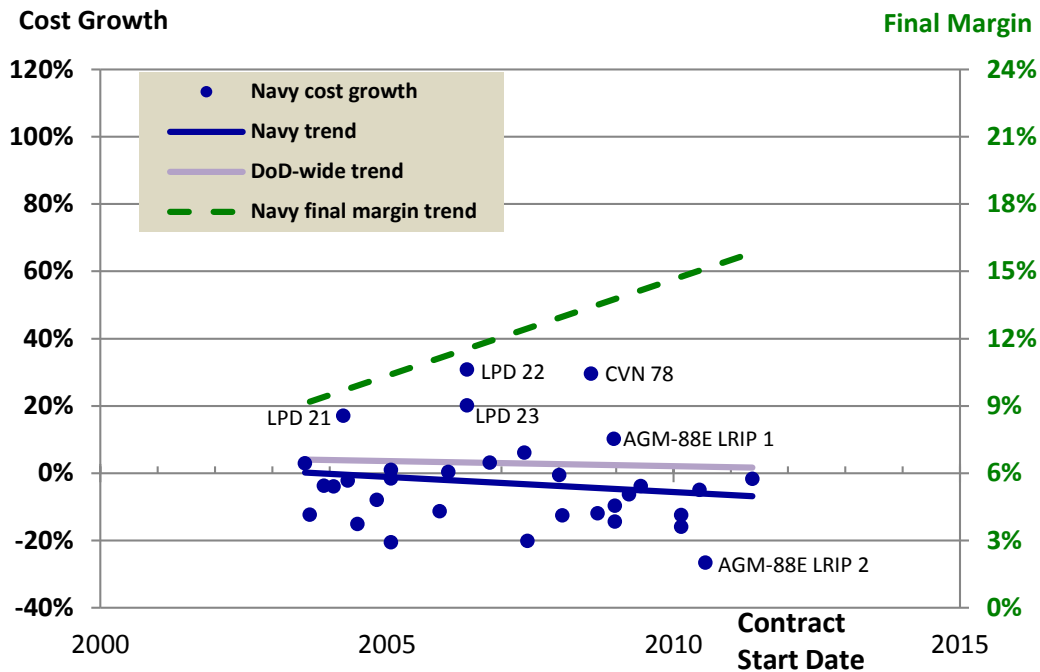
NOTE: Only the green dashed trend line is final margin, read from the right-hand scale. Individual data points and other trends are cost growth, read from the left-hand scale. The DoD trend is for all contracts and shown for comparison to the Navy trend during the same time period. Cost growth is in real terms (i.e., adjusted for inflation). The Navy trend is not statistically significant. Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Figure 2-45. Air Force Production Contract Cost-Growth and Final-Margin Trends (2004–2012)
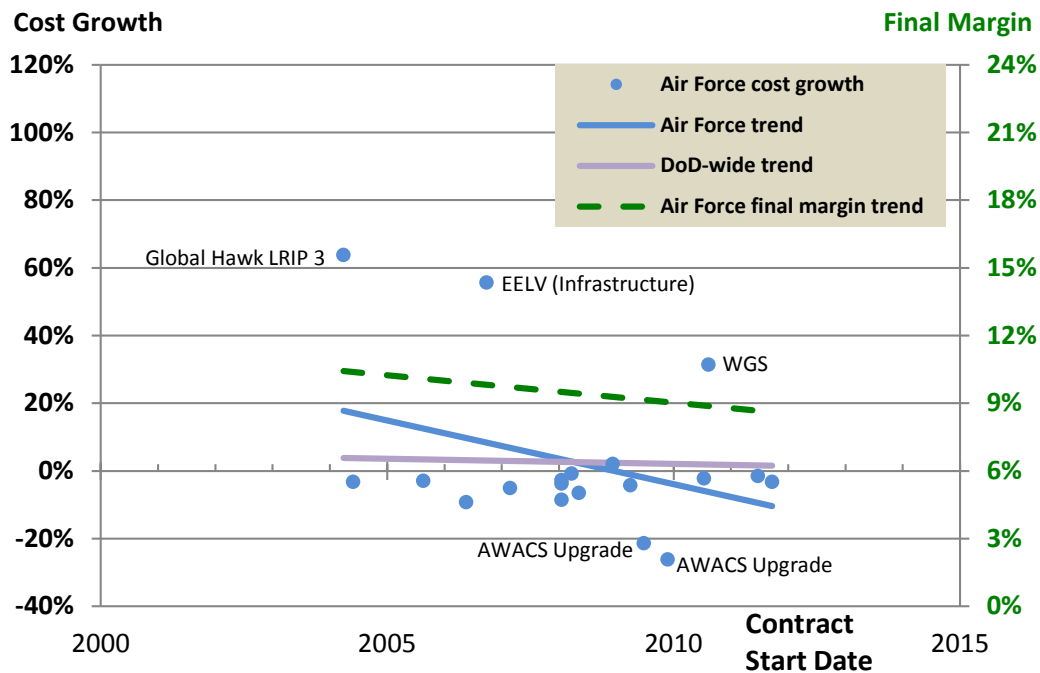
NOTE: Only the green dashed trend line is final margin, read from the right-hand scale. Individual data points and other trends are cost growth, read from the left-hand scale. The DoD trend is for all contracts and shown for comparison to the Air Force trend during the same time period. Cost growth is in real terms (i.e., adjusted for inflation). The Air Force trend is not

statistically significant. Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Contract-Level Schedule Growth by Military Department: Production

Finally, schedule growth in production is not clearly improving as it is in development. The Air Force has a statistically significant downward trend (see Figure 2-46), but the Navy is apparently flat (see Figure 2-47) and the Army apparently is increasing slightly (see Figure 2-48). Standard tests were used to eliminate any outliers that unduly distort the underlying trends shown (see Appendix A).

Figure 2-46. Army Production Contract Schedule-Growth Trends (2005–2010)
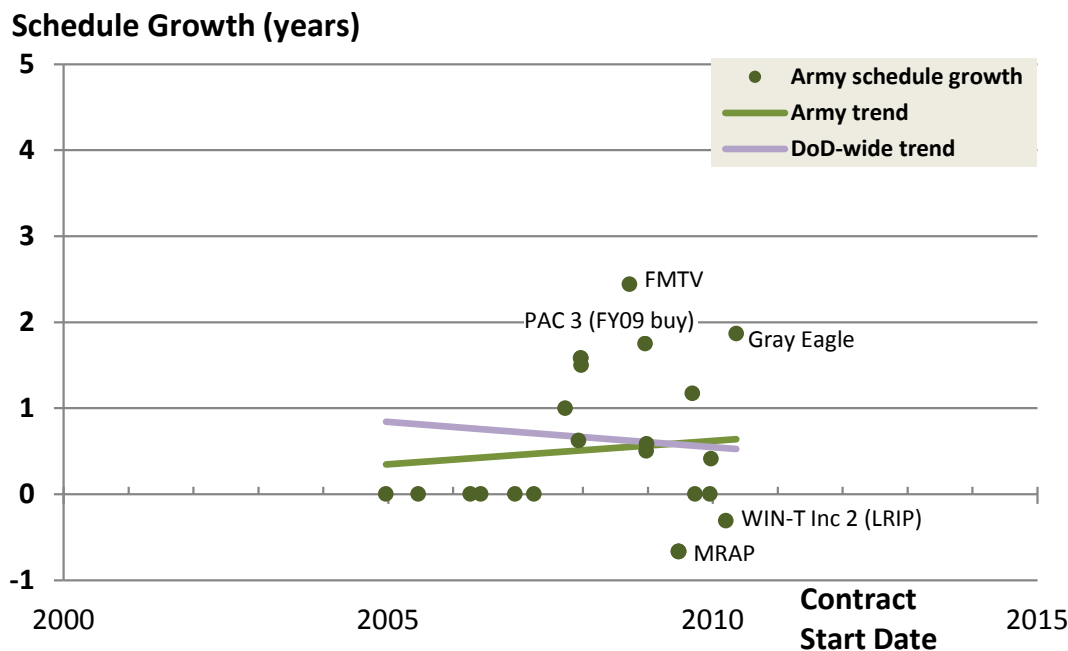
NOTE: The DoD trend is for all contracts and shown for comparison to the Army trend during the same time period. The Army trend is not statistically significant. Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Figure 2-47. Navy Production Contract Schedule-Growth Trends (2003–2012)
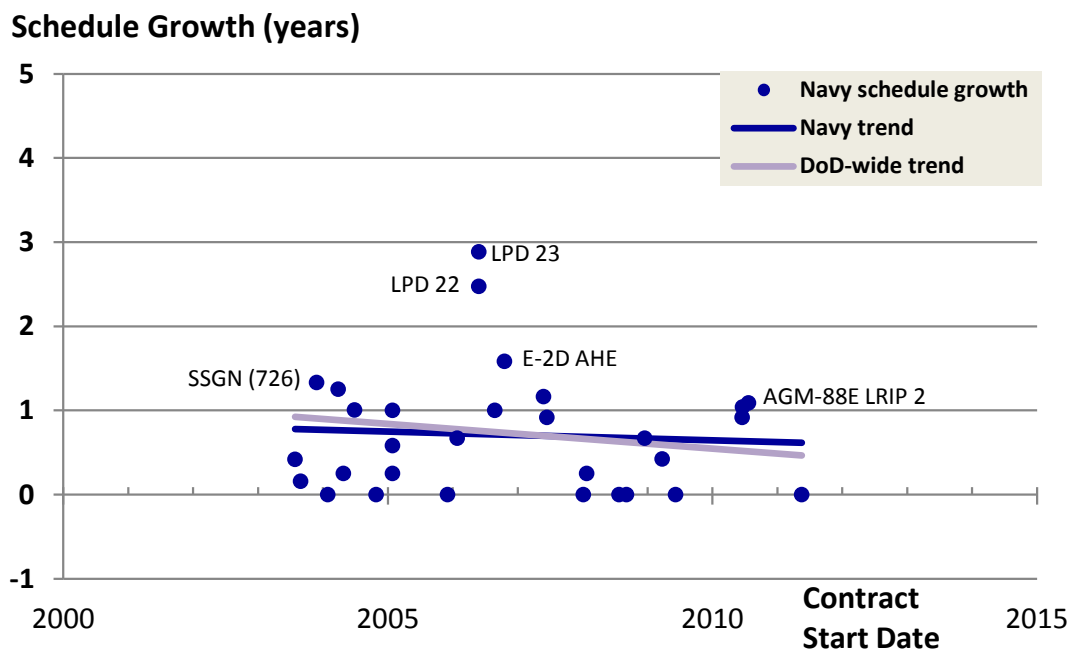
Schedule Growth (years)

NOTE: The DoD trend is for all contracts and shown for comparison to the Navy trend during the same time period. The Navy trend is not statistically significant. Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

Figure 2-48. Air Force Production Contract Schedule-Growth Trends (2004–2011)

Schedule Growth (years)

NOTE: The DoD trend is for all contracts and shown for comparison to the Air Force trend during the same time period. The Air Force trend is statistically significant. Statistical tests were used to eliminate outliers that unduly distort the underlying trend (see Appendix A).

INSTITUTIONAL ANALYSES: PRIME CONTRACTORS

The defense industry is a major part of the defense acquisition system, and most acquisition funds are spent on contractors. Below, we provide an overview of which major primes dominate specific portfolios of the goods and services we acquire in FY13. We then examine the performance of primes on MDAP contracts since 2000, followed by analysis of operating margins in the defense and commercial sectors.

Prime Contractors' Share of Spend

Figure 2-49 shows the largest five prime contractors (often referred to simply as "primes") based on dollars obligated overall and by products and services in FY13. Figure 2-50 provides more details by showing the largest five prime contractors by all DoD-obligated dollars in FY13 for each major portfolio grouping in products (supplies and equipment). Figure 2-51 shows the same for acquired services. Note that the largest primes are not always the same as they are DoD-wide. In other words, the largest prime vendors noticeably vary across product and service portfolios. On the right side of each figure we also show the fraction of contract dollars in each portfolio that are competed.

Figure 2-49. Largest Five Prime Contractors and Fraction Competed for Total DoD, Products, and Contracted Services (by dollars obligated in FY2013)
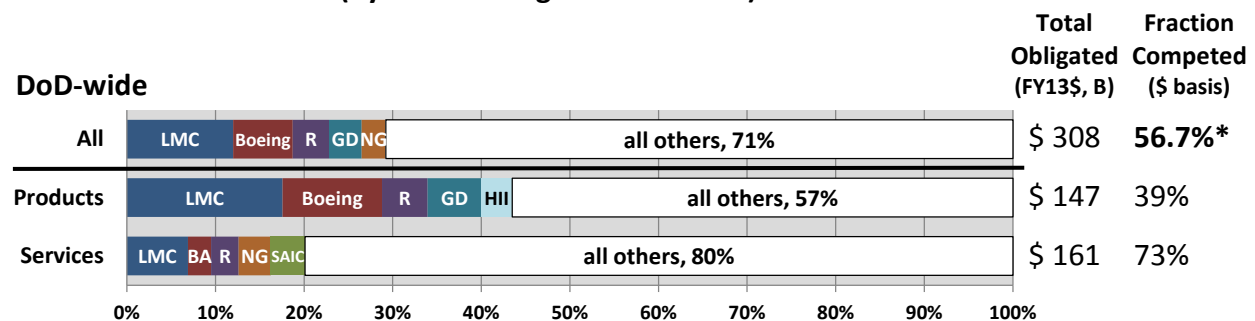
*overall goal was 60% for FY2013

NOTES: Total obligations include competed and not competed. Fractions of contracts competed are by dollars (not by number of awards). Competition goal is for FY2013. BA = Boeing; GD = General Dynamics; HII = Huntington Ingalls Industries; LMC = Lockheed Martin; NGC = Northrop Grumman; R = Raytheon.

Figure 2-50. Top Five Prime Contractors and Fraction Competed for Product Spend Portfolios (by dollars obligated in FY2013)
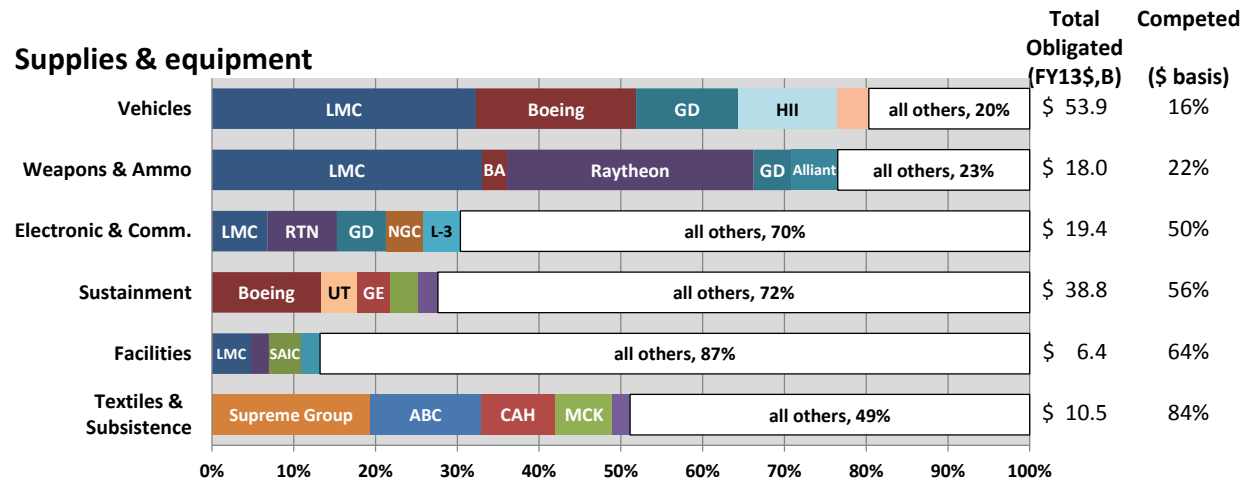
	Total Obligated (FY13$,B)	Competed ($ basis)
Vehicles	$ 53.9	16%
Weapons & Ammo	$ 18.0	22%
Electronic & Comm.	$ 19.4	50%
Sustainment	$ 38.8	56%
Facilities	$ 6.4	64%
Textiles & Subsistence	$ 10.5	84%

NOTES: Total obligations include competed and not competed. Fractions of contracts competed are by dollars (not by number of awards). ABC = AmerisourceBergen; BA = Boeing; CAH = Cardinal Health; GD = General Dynamics; HII = Huntington Ingalls Industries (includes legacy Northrop Grumman Shipbuilding contracts); LMC = Lockheed Martin; MCK = McKesson; NGC = Northrop Grumman; RTN = Raytheon; SAIC = Science Applications International Corporation; and UT = United Technologies. Obligations are in billions of FY13 dollars.

Figure 2-51. Top Five Prime Contractors and Fraction Competed for Contracted Service Spend Portfolios (by dollars obligated in FY2013)
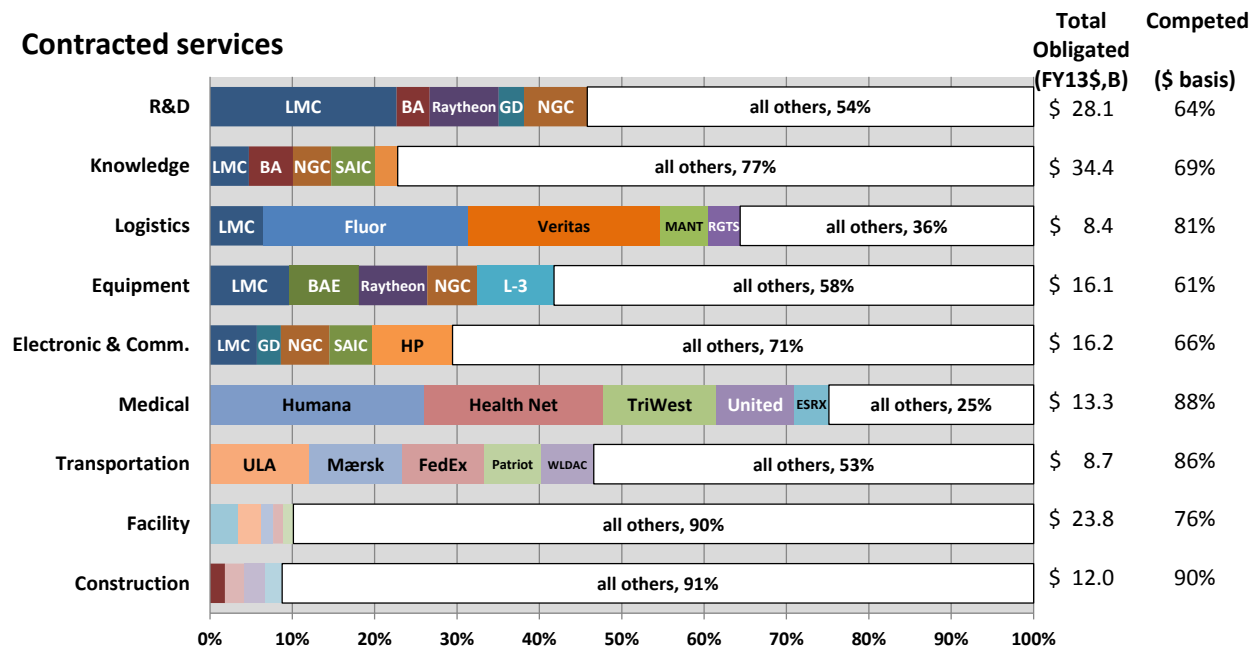
	Total Obligated (FY13$,B)	Competed ($ basis)
R&D	$ 28.1	64%
Knowledge	$ 34.4	69%
Logistics	$ 8.4	81%
Equipment	$ 16.1	61%
Electronic & Comm.	$ 16.2	66%
Medical	$ 13.3	88%
Transportation	$ 8.7	86%
Facility	$ 23.8	76%
Construction	$ 12.0	90%

NOTES: Total obligations include competed and not competed. Fractions of contracts competed are by dollars (not by number of awards). BA = Boeing; BAE = BAE Systems; ESRX = Express Scripts; GD = General Dynamics; HP = Hewlett Packard; LMC = Lockheed Martin; MANT = ManTech; NGC = Northrop Grumman; RGTS = Al Raha Group for Technical Services; SAIC = Science Applications International Corporation; ULA = United Launch Alliance; and WLDAC = World Airways. Obligations are in billions of FY13 dollars.

Recent Performance of Primes on Completed MDAP Contracts

To compare recent prime-contractor performance on major MDAP contracts, we examined price and schedule growth on a set of 83 development and 83 production contracts with start dates predominantly since the year 2000. All these contracts completed the vast majority of their deliverables, and we had readily available data on both initial (estimated at contract start) and final values, allowing us to calculate growth since inception.

Six contractors had the most development contracts in this dataset: Lockheed, Boeing, Northrop Grumman, General Dynamics, Raytheon, and BAE Systems. In production, eight contractors had the most contracts in this dataset: Lockheed Martin, Boeing, Northrop Grumman, General Dynamics, Raytheon, Huntington Ingalls, Oshkosh, and the United Launch Alliance (ULA). The remaining prime contractors in our datasets each had very few contracts in this dataset, so we grouped them in our analysis (see Table 2-11).

Table 2-11. Prime Contractors in the "Other Contractors" Groups

Development	Production
Alliant	BAE Systems, Inc.
Austral	General Atomics
Bell Textron	Alliant
Computer Sciences Corporation	Stewart and Stevenson
General Atomics	
General Electric	
Huntington Ingalls	
ITT	
Rockwell Collins	
SAIC	
Sikorsky	
Tybrin Electronics	
VIASAT	
VT Halter Marine	

Performance in Development

Figure 2-52 compares price growth and schedule growth performance of these prime contractors on 83 development contracts. Cost growth performance was similar to that shown for price growth. Growths are weighted by spend (i.e., are on a dollar basis rather than by contract), which provides a better view of institutional performance, given that these contracts vary greatly in size. Cost and price growth are adjusted for inflation.

Figure 2-52. Price and Schedule Growth by Prime Contractor in Development (2000–2013; weighted by spend)
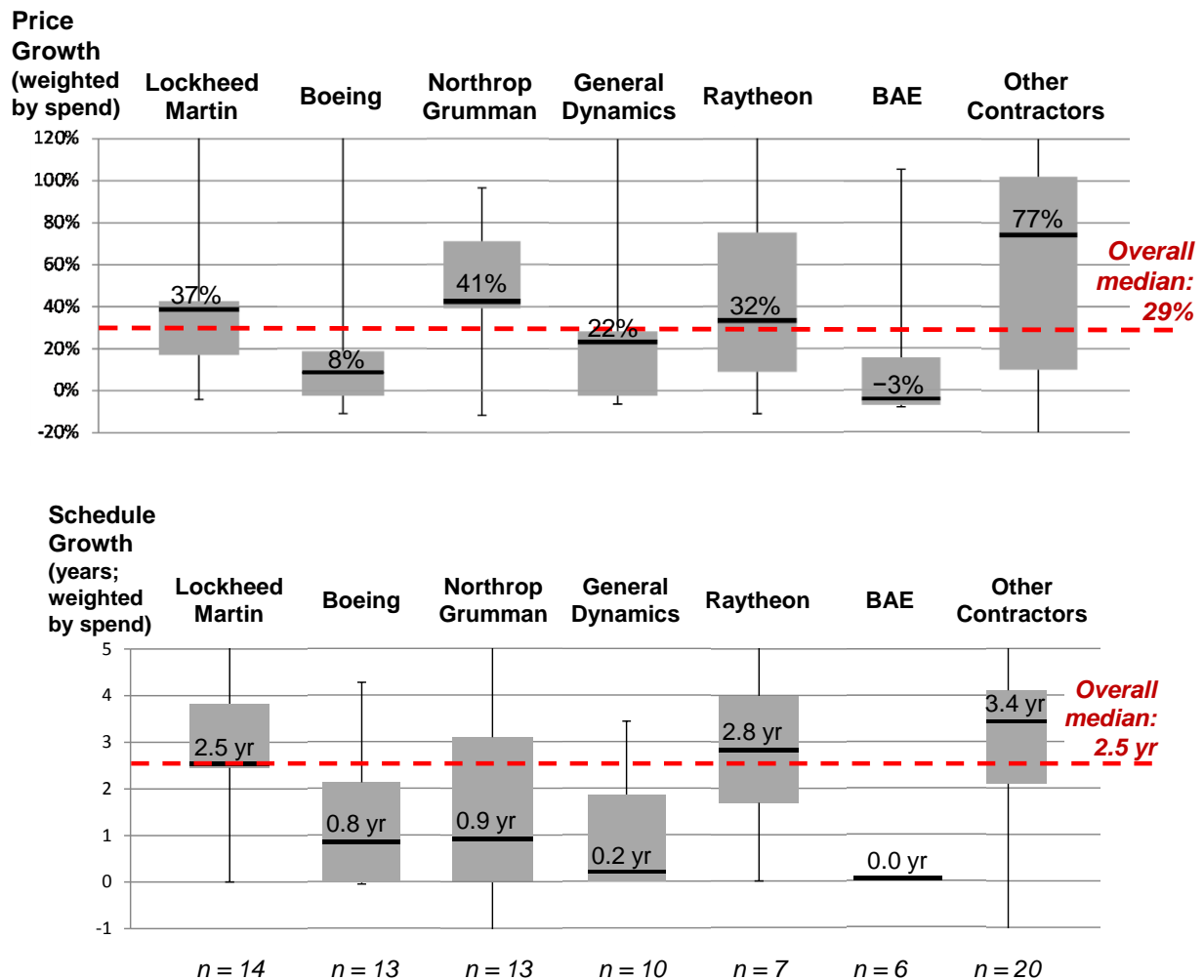
NOTE: Values are weighted by spend (i.e., are on a dollar basis instead of a contract basis). Growths are from contract initiation. Cost-growth distributions were similar to those shown here for price growth, and we adjusted for inflation on both cost and price growth. F-35 (Lockheed Martin) and KC-46A (Boeing) development contracts, while not completed, were included because they have executed enough to get an indication of their performance. Some individual contracts had particularly large representation by spend for some primes: F-35 (Lockheed Martin); NPOESS (Northrop Grumman); DDG 1000 (Raytheon); PIM (BAE); and Comanche (Other Contractors).

The median weighted cost growth was 41 percent, and the median weighted price growth was 29 percent. For both Northrop Grumman and the pool of remaining Other Contractors, cost and price growth were higher than the overall medians. Cost and price growth by BAE Systems and Boeing were much lower than overall; General Dynamics' were also lower to a lesser extent.

In terms of schedule growth, BAE Systems had the best performance (with no schedule growth on any of its six contracts), followed by General Dynamics, Boeing, and Northrop Grumman. The

overall median was 2.5 years. Again the group of Other Contractors performed the worst (as a group).

The cause of cost and price growth cannot be ascertained with this dataset because we do not have the CBB and thus do not have a direct measure of work content on each contract. Despite this limitation, both the government and the contractors play roles in identifying technical risks, estimating costs, scoping work, and avoiding premature starts. Thus, cost growth and price growth are relevant metrics, and differences between organizations indicate places for further investigation to understand the causes of systematic growth and identify remedies to minimize growth.

In terms of commodity types for high cost- and price-growth primes, 65 percent of the Other Contractors' spend was for helicopters, which exhibit high cost growth (e.g., see the earlier Table 2-5, which shows helicopters at the program level have a high Nunn-McCurdy breach rate). Northrop Grumman contracts by spend, on the other hand, were half for space and one-third for UAVs. The four primes that had the highest cost growth (Other Contractors, Northrop Grumman, Lockheed Martin, and Raytheon) all received two-thirds or more of their dollars through cost-plus-award-fee (CPAF) contracts. This type of contract is now discouraged for most purposes because it has not provided an effective incentive to control costs in most cases.

Performance in Production

Figure 2-53 compares the price-growth and schedule-growth performance of our prime contractors on 83 production contracts. Again, cost-growth performance was similar to that shown for price growth. Price growth was weighted by spend (i.e., are on a dollar basis rather than by contract), providing a better view of institutional performance given that these contracts vary greatly in value. Cost and price growth are adjusted for inflation.

As expected, cost, price, and schedule growth were lower generally in production than in development. The median overall cost growth was 11 percent, and the median overall price growth was essentially equivalent at 12 percent. In production, Boeing and Huntington Ingalls had significantly higher cost and price growth than the others (although the sample size was very low at only four contracts with Boeing).

Cost and price growth for Lockheed Martin, Northrop Grumman, General Dynamics, Raytheon, and the group of Other Contractors in production were lower than the overall median.

In terms of schedule growth, Oshkosh was the best performer (with all contracts completing on or sooner than originally scheduled—although there were only four Oshkosh contracts in this dataset). General Dynamics and Raytheon were below the overall median. Other Contractors, Northrop Grumman, Boeing (note low sample size), and ULA had schedule growth higher than the overall median. The overall median was half a year (again, much better than on development contracts overall).

In terms of commodity types (not shown), one striking differences was that General Dynamics and Huntington Ingalls contracts were both dominated by shipbuilding, but General Dynamics

had significantly lower cost, price, and schedule growth. This difference was probably not due to contract type since both were dominated by the same contract types.

As with the development analysis, the cause of cost and price growth on these production contracts cannot be ascertained with this data.

Figure 2-53. Price and Schedule Growth by Prime Contractor in Production (2000–2013; weighted by spend)
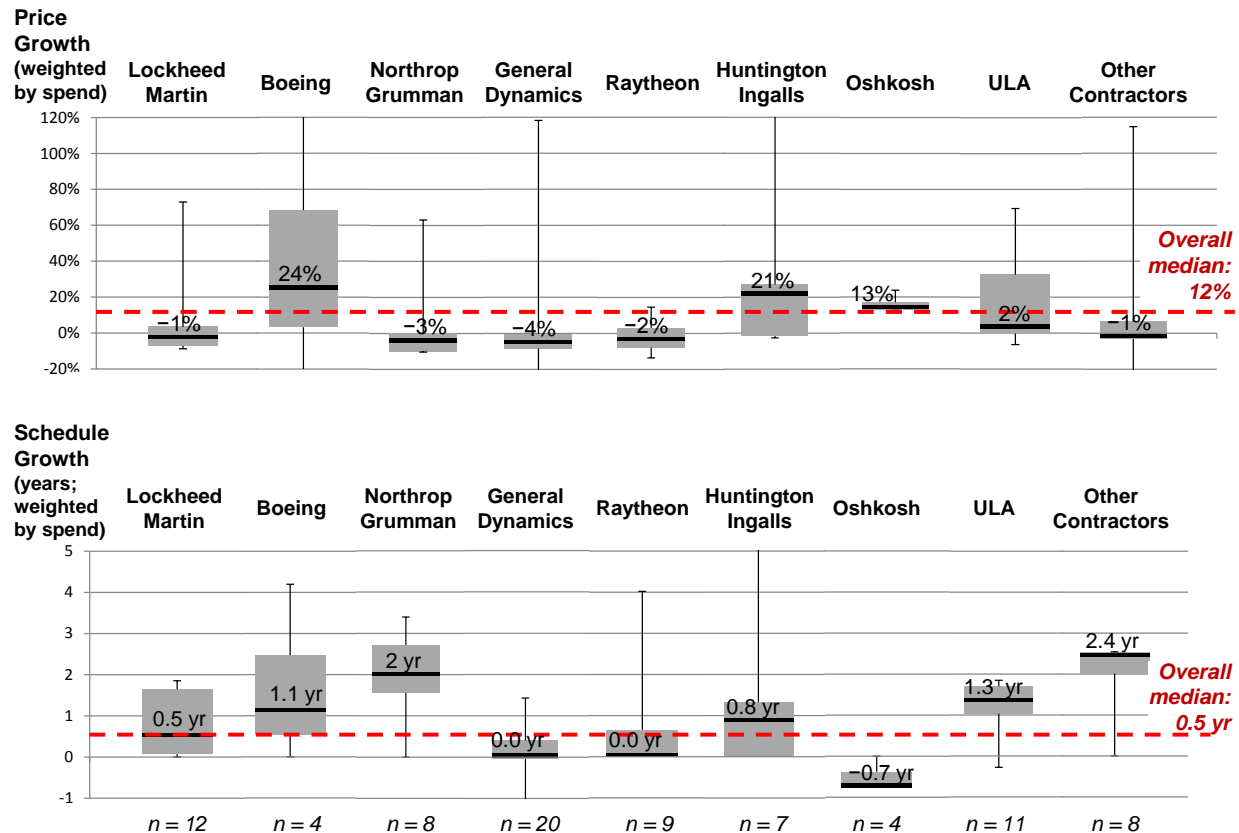
NOTE: Values are weighted by spend (i.e., are on a dollar basis instead of a contract basis). Growths are from contract initiation. Cost-growth distributions were similar to those shown here for price growth, and we adjusted for inflation on both cost and price growth. Some individual contracts had particularly large representation by spend for some primes: WGS (Boeing), CVN 78 (Huntington Ingalls), FMTV (Oshkosh), and FMTV (Other Contractors group). All ULA contracts were for EELV, dominated by two infrastructure contracts.

Operating Margins of Defense and Commercial Sectors

In examining defense contractor performance relative to margins, we examined the margins of defense firms relative to non-defense firms. We conducted a case study of data from public U.S. Securities and Exchange Commission (SEC) filings of defense- and commercial-sector companies or their associated divisions. Table 2-12 shows the operating margins reported by selected defense contractors in public SEC filings. These are compared to operating margins for selected

commercial firms (some of whom do business with the government but for which no more than 10 to 15 percent of their revenue is from sales to the government). The three categories of comparable commercial firms are producers of capital goods, engineering services, and automobiles and automotive parts. Data were provided by Bloomberg and tabulated by Capital Alpha Partners.

Table 2-12 shows that data from 2009 to 2013. The year 2009 was the first full year after the financial crisis and recession that started in 2008. In this particular year, there was no statistically significant difference between the margins of defense firms and firms in the capital goods industry. The operating margins of defense firms were on average higher than those earned by providers of engineering services and producers of automobiles and automotive parts. This difference was statistically significant. Furthermore, when comparing defense firms and non-defense firms as a whole in these data, the margins for defense in 2009 were higher than the margins of non-defense firms and the difference was statistically significant. So in 2009, defense firms weathered the crisis better than their commercial counterparts.

In subsequent years a pattern emerges. The defense firm's margins are lower that the margins for firms making capital goods. However, although this difference is statistically significant (i.e., we can confidently assess that the difference is not zero), the average margins are comparable between these two groups.

On the other hand, margins for defense firms are larger than margins for firms providing engineering services—and the difference is statistically significant. With the exception of 2009 and 2010 (when defense margins were systematically larger than margins earned by firms producing automobiles and automotive parts), defense margins were not statistically distinguishable from margins in the automobile and automotive-parts sector.

Overall, from 2010, the margins earned by defense firms were not systematical different from the margins of commercial firms shown as a whole. Furthermore, the 9 percent average operating margin for defense firms reported in filings is quite consistent with the average margin of 8.4 percent in our sample of 166 MDAP development and production prime contracts taken as a whole. It is clear that defense firms generally are earning adequate margins, that on the whole do not vary significantly year to year. Within these results, there is adequate opportunity to provide effective incentives to industry without changing aggregate returns for defense firms in general.

Table 2-12. Operating Margins of Some Defense- and Commercial-Sector Firms (2009–2013)

Averages (percent)	2009	2010	2011	2012	2013
Defense	9	9	9	9	9
Commercial	* 6	10	11	10	11
Capital goods	10	‡ 13	‡ 14	‡ 14	‡ 14
Engineering services	* 5	* 5	* 6	* 5	* 6
Automobiles and automotive parts	* −3	* 7	8	6	7
Defense					
Boeing *(BDS only)*	10	13	10	9	10
CACI	7	6	7	8	7
General Dynamics *(except Aerospace)*	11	11	12	8	8
Huntington Ingalls	3	4	2	5	7
ITT Defense/Exelis	12	11	8	11	10
L-3 Communications	11	11	11	10	10
Lockheed Martin	10	9	9	9	11
Mantech	9	8	8	7	6
Northrop Grumman	8	10	12	12	13
Raytheon	12	10	11	12	12
Capital goods					
Caterpillar	2	9	12	14	9
Cummins	6	12	15	11	12
Danaher	14	16	16	17	17
Dover	10	15	15	16	16
Eaton	4	9	10	9	12
Emerson Electric	13	15	16	18	17
Flowserve	14	14	14	14	15
Honeywell	13	14	15	11	16
Illinois Tool Works	10	15	15	17	18
Ingersoll-Rand	7	9	10	9	10
Joy Global	20	20	21	21	19
Lennox	4	6	4	7	9
Parker Hannifin	8	9	12	13	11
Rockwell Automation	8	13	15	16	17
Engineering services					
Aecom	4	5	5	4	5
CBI	7	8	8	8	7
Fluor	5	3	4	3	4
Jacobs Engineering	5	4	5	5	6
Automobiles and automotive parts					
Borg Warner	1	9	11	10	12
Delphi		7	10	10	11
Ford	−1	7	6	4	4
General Motors	−20	4	5	−2	5
Johnson Controls	1	5	5	4	4
TRW	4	9	8	8	7

NOTE: *****Statistically lower than defense margins. **‡**Statistically higher than defense margins. Commercial average is that for the 23–24 margins shown for all three sectors. SOURCE: Public SEC filings (Bloomberg, Capital Alpha Partners).

(This page was left blank intentionally)

3. EFFECTIVENESS OF ACQUISITION POLICIES AND TECHNIQUES

ON INCENTIVES

Performance is influenced by incentives. How well any organization performs depends on its incentives. Thus, a key element for improving acquisition performance is improving how contract incentives are aligned with our performance objectives, and how effective those incentives are when measured against those performance objectives. Without effective alignment our contractors will not make their best effort to deliver the quality products and services are warfighters and taxpayers expect and deserve..

We have a wide range of incentive structures available for motivating contractor performance, including:

- Incentive fees tied to performance objectives of importance to the government
- Award fees tied to subjective measures of performance
- Execution of options for continued work in lieu of competition when in the interest of the government
- Payments tied to specific performance objectives
- Event-based contract obligations tied to successful completion of work scope

We examined trends in how effective the DoD has been in the last decade at using margin/markup to motivate cost and schedule performance—measured directly or through contract types.

Industry Performance Measures

Effectively motivating industry requires an understanding of industry's compensation system. To understand how effective these incentives may be for motivating industry performance, let us review briefly the basis for executive compensation such as cash bonuses, stock options, and stock. Thompson et al. (2013) identified the following major performance measures for public and private companies (generally in priority order in their latest survey):

- Company, department, and individual goals and objectives
- Discretionary
- Earnings (e.g., earnings before interest and taxes [EBIT] and also before depreciation and amortization [EBITDA])
- Cash flow
- Net income
- Share and stock price
- Revenue growth

- Performance against companies with a peer group
- Earnings per share
- Return on assets, capital, or equity
- Economic value added

Many of these relate directly to profit or fee, but not all. Cash flow, for example, is prominent (although declining in importance in their surveys since 2009). Stock price is listed lower in terms of direct measure on executive compensation, although stock options can be much more valuable than salaries and bonuses. Still others are not specified in this survey (e.g., company, department, and individual goals and objectives).

Thus, we fully expect profit and fee earned on contracts to have some effect on motivating industry performance, but we recognize other market incentives are also at work.

EFFECTIVENESS OF FINAL PROFIT OR FEE AS INCENTIVES FOR CONTRACT PERFORMANCE

We examined whether profit or fee has been effective in motivating better performance on DoD contracts.

New analysis of MDAP contracts with prime contractors in the last decade found that realized profit and fee levels—expressed as margin percentages of price or cost, respectively—were *not* generally contingent on performance; cost, price, and schedule growth were not linked statistically to final margin in this sample.

In development, Table 3-1 shows the results of a multivariate regression to determine what variables correlate with final contract margin. The only significant variables were contract size (as measured by schedule) and margin change from initial expectations. On a median contract DoD-wide, the regression predicts a margin decrease of almost 1 percent since contract initiation and an increase of about 1.6 percent due to median size (measured by schedule) on top of an approximately 6 percent constant. The other variables that are listed at the bottom of the table were tested but found to be insignificant relative to margin.

Key point: The major observation from the results in Table 3-1 is *not* what the final predicted margin is from this dataset but that margin did not systematically vary by cost or schedule performance. In other words, across all the production contracts in this dataset, *we were not systematically adjusting final margins based on cost or schedule performance*. From this incentive perspective, margin acceptability (i.e., whether 6 to 7 percent is acceptable in absolute value) is less important than whether we are adjusting margin to reward good performance and penalize bad performance. This finding is not to say there are not subsets of the data where we did adjust margins based on cost performance. For that, we examined the effects of contract types that contain an explicit, focused incentive for controlling cost growth.

Table 3-1. Factors That Affected Final Margin/Markup on Development Contracts (2000–2012)

Significant variables	Expected effect on DoD-wide final value (at the <u>median</u>)	
	Margin	*Markup*
Size (measured by schedule)	1.6 %	1.6 %
Margin/markup change since initiation	−0.9 %	−0.8 %
Exogenous constant	+5.7 %	+6.0 %
Predicted final	6.4 %	6.8 %

Insignificant variables
(*no* statistically significant effects on final margin/markup)

- Cost growth
- Price growth
- Schedule slips
- Army, Navy, or Air Force
- Contract type
- Quantity change
- Commodity type (9 types)
- Overhead share of costs
- Fixed-cost share of costs
- Time trend
- Size of contract by spend

Controlling for the two variables above, final margin/markup was not *predicted by cost, price, or schedule performance.*

NOTE: Regression over 81 MDAP development contracts, 95 percent of which were started in or after 2000. Price and cost growth are adjusted for inflation. Margin is on a price basis, whereas markup is on a cost basis.

In production, we found a similar result. The multivariate regression summarized in Table 3-2 found no statistical correlation of cost, price, or schedule growth across all production contracts with margin/markup. In addition to margin/markup change, three new significant variables that correlated with final contract margin/markup were whether the contract was for an LRIP, ship, or space program (or not). For a median contract DoD-wide, LRIPs, ships, and space contract each lowered the expected median by 1 percent. Unlike in development, margin changes increased about 1 percent at the median on top of a 14 percent constant (16 percent markup), yielding a net expected margin of about 11 percent (13 percent markup) for this population. The other variables insignificant relative to margin/markup are listed at the bottom of the table.

Key point: In production (as in development), the major observation from the results in Table 3-2 is *not* what the final predicted margin/markup is from this dataset, but that margin/markup did not systematically vary by cost or schedule performance. In other words, across all the

production contracts in this dataset, *we were not adjusting final margin/markup based on cost or schedule performance*. From this incentive perspective, assessing the acceptability of a particular margin or markup level in production (i.e., whether 11 to 12 percent is acceptable or not in absolute value) is less important than whether we are rewarding good performance and penalizing bad performance. As in development, we found certain subsets of production contract types did show cost growth affecting final margin/markup.

Table 3-2. Factors That Affected Final Markup and Margin on Production Contracts (2000–2012)

Significant variables	Expected effect on DoD-wide final value (at the <u>median</u>)	
	Margin	*Markup*
LRIP	−2 %	−2 %
Ship	−1 %	−1 %
Space	−1 %	−1 %
Margin/markup change since initiation	+ 1 %	+ 1 %
Exogenous constant	+14 %	+16 %
Predicted final	11 %	13 %

Insignificant variables
(*no* statistical significant effects on final margin/markup)

- Cost growth
- Price growth
- Schedule slips
- Army, Navy, or Air Force
- Contract type
- Quantity change
- Other commodity type (7)
- Overhead share of costs
- Fixed-cost share of costs
- Time trend
- Size of contract by spend
- Size of contract by schedule

Controlling for the two variables above, final margin/markup was not *predicted by cost, price, or schedule performance.*

NOTE: Regression over 81 MDAP production contracts, 95 percent of which were started in or after 2000. Price and cost growth are adjusted for inflation. Margin is on a price basis, whereas markup is on a cost basis.

EFFECTIVENESS OF PRODUCTION PROFITS AS INCENTIVES FOR DEVELOPMENT SCHEDULE PERFORMANCE

A different use of contract margin as incentives involves offering the prospect of higher levels on future contracts to motivate better schedule performance on current contracts. Final contract margins in production are higher generally than those in development (see Figure 3-1) and this fact should serve as a strong motivation to complete development successfully on time and transition to production.

We also regressed development-schedule-growth effect on final production margin for the same program and found a correlation (see Figure 3-2). Shorter schedule slips in development were rewarded with higher margins in production in this sample. This relationship is statistically significant although the sample is small. For every year of slip in development schedule, the regression model predicts a 0.5 percent drop in production margin. In other words, for a median development contract schedule slip, the regression predicts a 0.6 percent drop in production margin on top of a 12 percent constant.

While caution is warranted given the small sample size, this result is stronger than a theoretical postulate that higher production margins should motivate contractors to get out of development as soon as possible. Here we only examined contract pairs for the same program to reduce the effects from different commodities and programs. Moreover, this result actually shows that margins are lowered *on the same program* when development is longer. It is unclear whether these margin changes are consciously made by program managers to motivate contractor schedule performance or if this is a side effect of, say, program schedule growth on programs or budgetary scrutiny of delayed programs. It is also unclear whether contractors are aware that margins may be lower in production if they have development-schedule growth. We have considered making this linkage part of the development contract on a pilot basis, but we have not pursued it as a standard practice.

Note that higher production margins may lead to other consequences. Together with much lower competition in production and the promise of a long, relatively stable series of production contracts, this likely motivates contractors to "win the franchise" to become the sole-source producer of a system. Contractors are often willing to accept lower development margins to win this production franchise.

Figure 3-1. Final Margin on MDAP Development and Production Contracts (2000–2013)
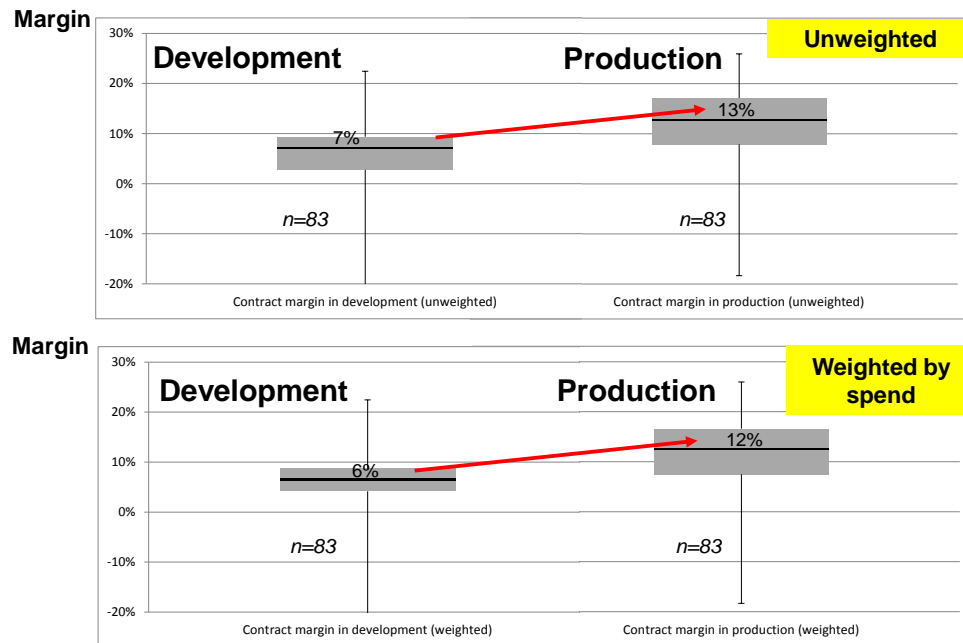
NOTE: Differences are substantial in both cases. The difference is also statistically significant when not weighted by spend. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Figure 3-2. Effect of Schedule Slip During Development on Final Margin in Production (2000–2013)

◆ Final margin in production ——— Fitted trend
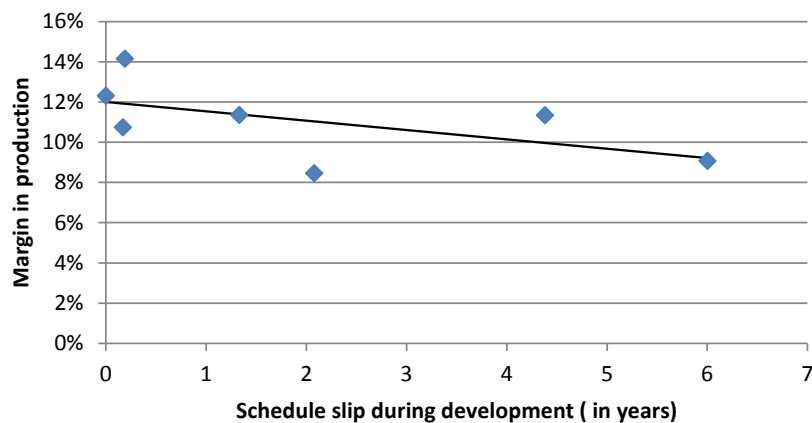
Predictor of final margin in production	Estimated coefficient (Standard error)	Significant?
Schedule slip during development (in years)	- 0.00465 (0.00207)	Yes (at 10% level of significance)
Constant	0.120 (0.00882)	Yes
Model R²	0.32	

n = 7

NOTE: Sample includes seven pairs of prime contracts from the period 2000–2013, where each pair is for the same MDAP.

EFFECT OF CONTRACT TYPES ON COST AND SCHEDULE CONTROL

Policymakers have long debated what types of contract vehicles should be preferred in acquisition and have experimented with blanket mandates. Current law mandates authorization and documentation by the MDA if cost-type contracts are used in MDAP development (see Section 818 of Public Law 109-364, the FY2007 NDAA). Also, the law prohibits cost-type contracts in MDAP production (see Section 811 of Public Law 112-239, the FY2013 NDAA). Our view is that contract types must be tailored to the specific situation, with the risk associated with the work to be performed being the most important factor.

Regression analysis provided in last year's performance report showed no statistical correlation in either development or early production between performance (cost or schedule growth) and broad contract type (fixed-price and cost-plus) when examining 433 MDAP development contracts and 440 early production contracts from 1970 to 2011.[16]

The following new analysis demonstrates why this is true and, more important, indicates that this simple bifurcation of contract types (grouping all cost-type contracts together and all fixed-price together) is misguided in the first place. A more informed partitioning is discussed below.

Cost Type Versus Fixed Price

Fixed-price contracts exhibit lower cost growth because they are used primarily in lower-risk situations—not because they inherently lead to lower cost growth. When combining contracts across phases and comparing their cost performance, it is true that the fixed-price set of contracts have lower cost growth. However, that does not mean that this lower cost growth is due to the fixed-price nature of the contract. *Fixed-price contracts are only fixed if the contractual work content and deliverables are fixed; they can be (and often are) modified to handle realized risks, leading to cost growth.* Instead, these fixed-price contracts have lower cost growth because they were used primarily in lower-risk cases (especially in production, where risks are lower than in development). Note that this statement is about cost, not about price. Price can be higher on a fixed price contract because the contractor prices in as much risk as possible in an attempt to shift risk to the DoD.

Figure 3-3 compares the cost growth of 166 MDAP contracts since 2000 in real terms (i.e., adjusted for inflation). Indeed, the cost growth for the FP contracts was statistically lower than that for all CP/H contracts.

This by itself is misleading, however, because fixed-price contracts dominate production wherein the cost growth risks are lower. Figure 3-4 shows that, when weighted by spend, only 10 percent of the development contracts in this dataset are fixed-price, whereas 73 percent of

---

[16]See USD(AT&L), 2013, pp. 51–54 and pp. 69–72.

the production contracts are fixed.[17] The cost growths by type (bottom plot) are similar to the cost growth by phase (top plot). Thus, fixed-price contracts exhibit lower cost growth *because* they are used primarily in production as well as in a small number of lower-risk development cases.

Figure 3-3. Comparison of Contract Cost Growth by Contract Type (all contracts together, 2000–2013)
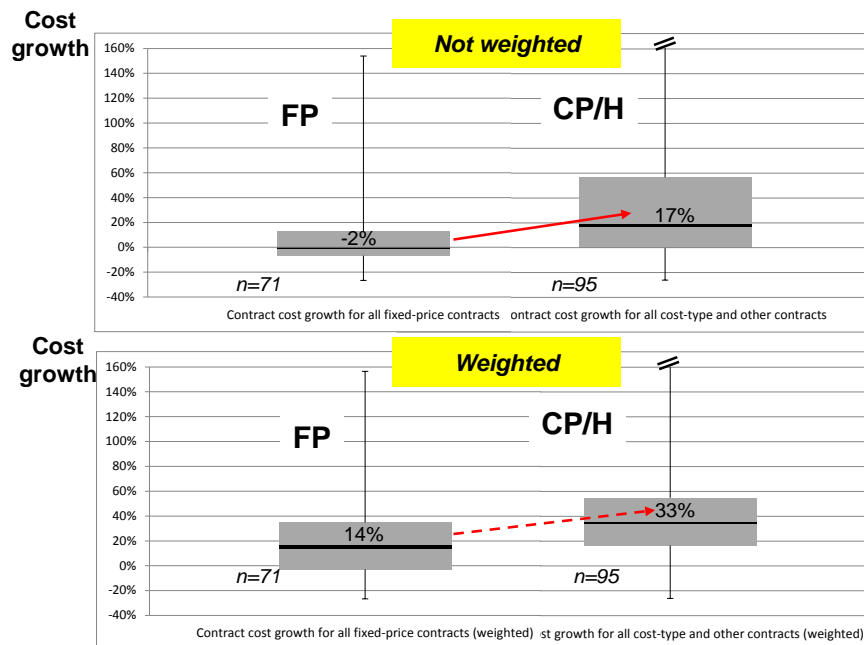
NOTE: Cost growth is adjusted for inflation. "FP" are fixed-price contracts (e.g., FPIF and FFP), and "CP/H" are cost-type and hybrid contracts (e.g., cost-plus-fixed-fee [CPFF], CPAF, cost-plus-incentive-fee [CPIF], cost sharing [CS], and hybrid). Weighting in the bottom plot is by spend. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

---

[17]This distribution provides evidence that (at least for these prime MDAP contracts since 2000), the DoD contracting community is using the appropriate contract type for phase and risk as per guidance (i.e., using fixed-price contracts for lower-risk cases). Because fixed-price contracts can still be modified if necessary to accommodate engineering problems, more contractual flexibility through cost-type contracts in higher-risk developmental stages is still an appropriate consideration.

Figure 3-4. Comparison of Contract Cost Growth by Phase and by Contract Type (2000–2013)
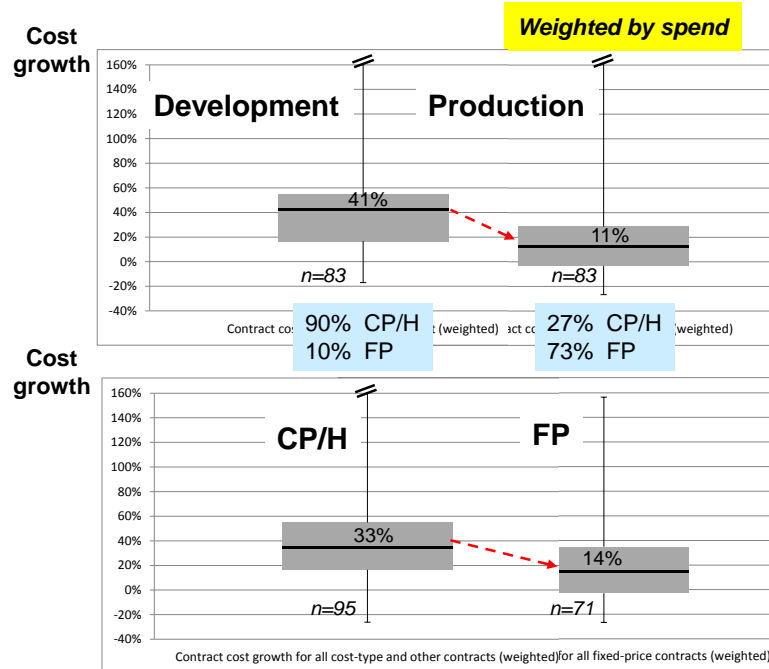
NOTE: Cost growth is adjusted for inflation. "FP" are fixed-price contracts (e.g., FPIF and FFP), and "CP/H" are cost-type and hybrid contracts (e.g., CPFF, CPAF, CPIF, CS, and hybrid). Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Predetermined, Formula-Type Incentive Contracts

If fixed-price contracting does not necessarily lead to better cost control, what does?

When cost control is predetermined and formulaically incentivized in the contract, vendors respond. The key is predictable incentives, not fixed pricing. This may seem obvious in hindsight, but simply considering fixed-price contracts is too simplistic because (1) some cost-type contracts[18] also have strong incentives to control cost growth, (2) both cost-type and fixed-price contracts can be modified, and (3) some fixed-price contracts[19] do not provide information on costs. Without good government understanding of actual cost, the government risks paying unnecessarily high prices on FFP contracts. Also, FFP contracts do not share savings with the government—even when risks are low. Moreover, using FFP contracts in full-rate production may incentivize contractors to withhold cost-reduction ideas and investments until then, when they keep all the cost-reduction rewards. If the government does not have insight into these savings, subsequent production lot prices may continue to be higher than necessary.

There are two dominant contracting mechanisms in our MDAP data sample in which cost savings and cost increases are shared formulaically between the government and the

---

[18]Especially CPIF and CS contracts.

[19]Namely, FFP contracts.

contractor over the range of incentive effectiveness: CPIF and FPIF (see Federal Acquisition Regulation [FAR], Part 16). In these contracts the government and contractor share in overruns or underruns (usually up to a ceiling) above or below a specific amount (the target price) based on a formula, usually a percentage split at a fixed ratio. Cost Sharing (CS) contracts also contain a sharing formula and are used, for example, in prototyping, but they are rare in our sample. Also there are hybrid contracts for which major line items may be one or more of the three primary types listed above.

CPAF contracts do include some type of cost incentives, but their effectiveness is reduced by tying part of the awards to other factors (e.g., subjective assessments of performance). Moreover, since the formulas for awarding fees can be subjective and changed during the course of the contract, the effectiveness of the incentive in CPAF contracts is often further reduced. This is probably why the FAR states that award-fee contracts should be used when it is "neither feasible nor effective to devise predetermined objective incentive targets applicable to cost, schedule, and technical performance" (FAR, Subpart 16.401[e][1][i]).

To examine the effectiveness of predetermined formula-type incentive contracts, we compared the performance of CPIF and FPIF contracts against the others. Hybrid contracts with at least one major line item that was either CPIF or FPIF were included in the CPIF/FPIF set. While FFP contracts also have an explicit cost incentive, we examined them separately because the contractor keeps all the cost savings. Any remaining hybrid contracts with a major FFP line item were included in the FFP set. Note that there likely has been a significant selection bias against FFP in development and bias for FFP in full-rate production based on historically accepted best practices and guidance provided to the acquisition workforce prior to the Better Buying Power initiatives.

In development, over half of the contracts by count in our sample were neither incentive formula-type nor FFP contracts. Figure 3-5 compares the cost growth and final margin (weighted by spend) on MDAP development contracts since 2000. Incentive formula-type contracts had lower cost growth at the median (i.e., its central tendency), shared savings with the government, and earned slightly lower margins than CPAF/CPFF/CS/other (O) contracts. In other words, to achieve about the same margin on incentive formula-type contracts, contractors had to control cost growth better—and they did so. Note that there are very few FFP contracts in our development dataset due to selection bias away from the use of FFP in development; the few we did have showed larger cost growth and lower margins as a result. Also, the distributions for price growth were similar to those for cost growth in development (see Table 3-3).[20]

Similarly in LRIP, incentive formula-type contracts had much lower cost growth despite providing similar final margins as CPAF/CPFF/CS/O contracts and sharing cost savings with the government (see Figure 3-6). Note, however, that the sample size was very low for

---

[20]Recall that FFP contracts can have price growth if work content is changed through contract modifications.

CPAF/CPFF/CS/O contracts. Also, compared to the FFP contracts, the incentive formula-type contracts had better cost control despite sharing savings with the government and providing lower final margins to the contractors. The distributions for price growth were similar to those for cost growth in the LRIP phase (see Table 3-3).

In FRP, where contracts were overwhelmingly FFP, cost growth for these FFP contracts was the lowest (−4 percent at the median) of all contract types across all three phases (see Table 3-3). The two incentive formula-type contracts in our FRP dataset both had large cost growth (20 and 31 percent), but the price growth (zero and −3 percent, respectively) was negated by relatively large negative final margins (−18 and −5 percent, respectively). While caution is warranted given the very low sample count for formula-type contracts in FRP, these two results show that the incentive structure can work in FRP to control price risk for the government.

Figure 3-5. Price Growth and Final Margin in Development on MDAPs by Contract Type (2000–2013)
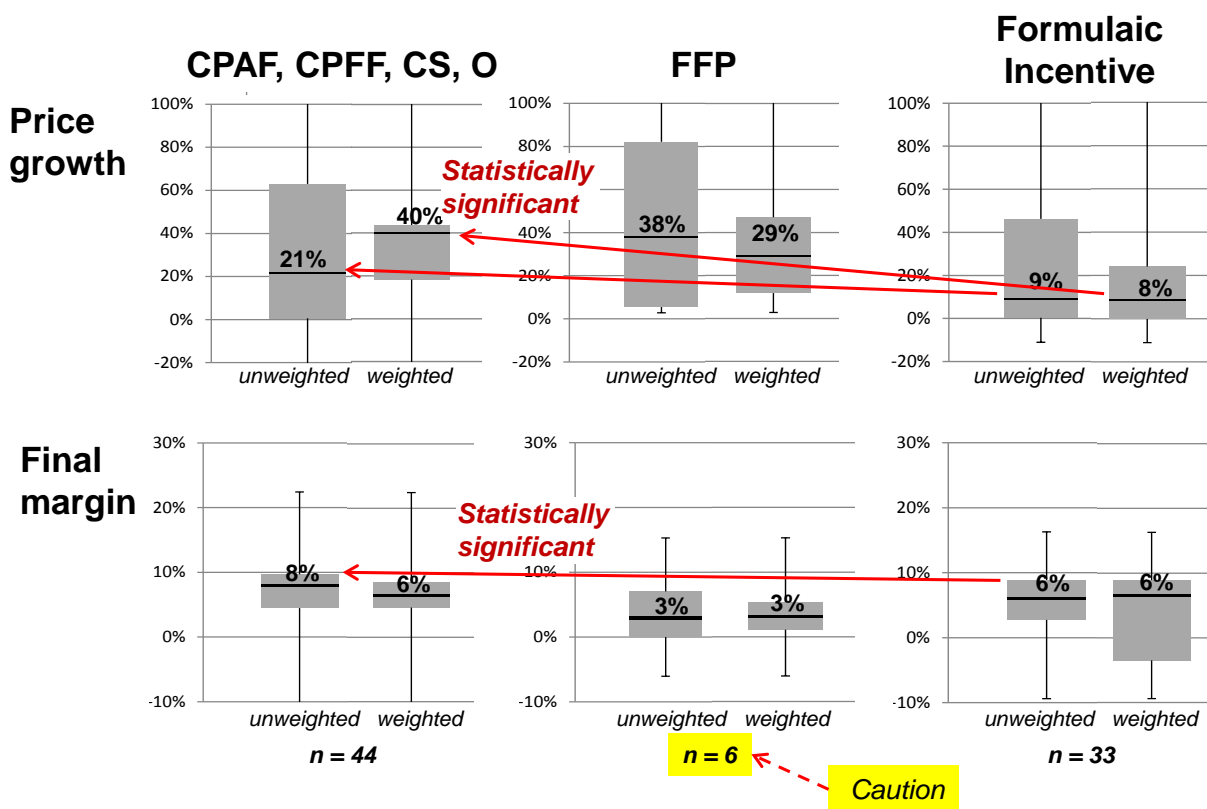
NOTE: There is a very small sample size for FFP development contracts (*n*=6). Cost growth is adjusted for inflation. Results are on a contract basis (i.e., not weighted by the dollar size of each contract). Price-growth distributions were similar to cost-growth distributions. Predetermined, formula-type incentive contracts are CPIF, FPIF, and hybrid contracts with a major CPIF or FPIF line item. "FFP" includes hybrid contracts with no major CPIF or FPIF line item, but with a major FFP line item. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Figure 3-6. Price Growth and Final Margin in LRIP on MDAPs by Contract Type (MDAP contracts, 2000–2013)
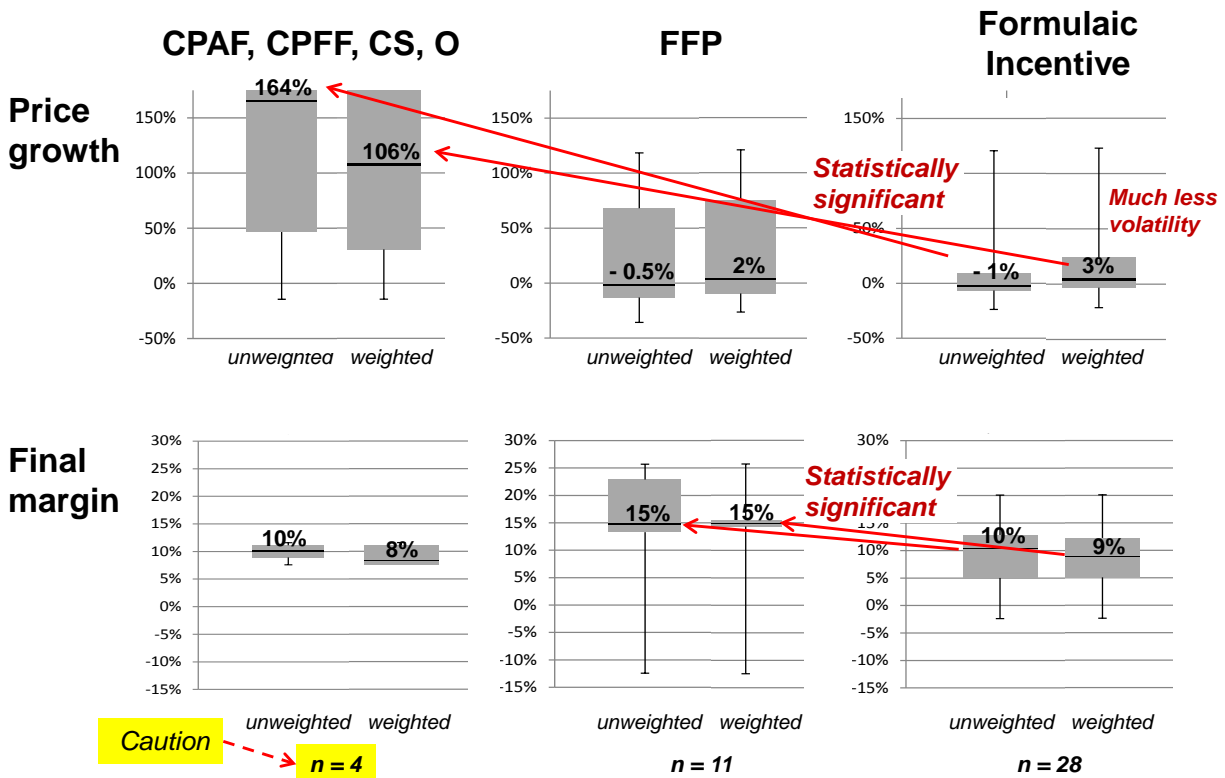
NOTE: There is a very small sample size for nonincentive LRIP contracts (*n*=4). Cost growth is adjusted for inflation. Results are on a contract basis (i.e., not weighted by the dollar size of each contract). Price-growth distributions were similar to cost-growth distributions. Predetermined, formula-type incentive contracts are CPIF, FPIF, and hybrid contracts with a major CPIF or FPIF line item. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Table 3-3. Cost and Price Growth Across Phases and Contract Types on MDAPs (2000–2012)

Cost Growth

Medians	CPAF, CPFF, CS, O		FFP		Formulaic Incentive	
Development	**27 %**	*(n=44)*	**39 %**	*(n=6)*	**14 %**	*(n=33)*
LRIP	**143 %**	*(n=4)*	**−0.3 %**	*(n=11)*	**−2 %**	*(n=28)*
Full-Rate Production	*none*		**−4 %**	*(n=38)*	**25 %**	*(n=2)*

Price Growth

Medians	CPAF, CPFF, CS, O		FFP		Formulaic Incentive	
Development	**21 %**	*(n=44)*	**38 %**	*(n=6)*	**9 %**	*(n=33)*
LRIP	**164 %**	*(n=4)*	**−0.5 %**	*(n=11)*	**−1 %**	*(n=28)*
Full-Rate Production	*none*		**−2 %**	*(n=38)*	**−1 %**	*(n=2)*

NOTE: No CPAF/CPFF/CS/Other contracts and only two formulaic incentive contracts are in our dataset for full-rate production. Cost and price growth are adjusted for inflation. Price growth distributions were similar to cost growth distributions. "Incentivized" contracts are CPIF, FPIF, and hybrid contracts with a major CPIF or FPIF line item.

Types of Contract Vehicles Employed

We examined the distribution of contract types we employ. Figure 3-8 plots the percentages by obligated dollars of contract types for all DoD contracts in FY13.[21] Figure 3-8 shows the breakdown by portfolio group.

- FFP contracts constitute about half of all DoD obligated contract dollars, including half for each of total products and services. FFPs are often used when purchasing goods and services from the broader commercial market.
- CPFF is used for over a quarter of the total obligations for services, but less so for procurement of supplies and equipment (goods). CPFF constitutes almost half of R&D and over half of medical services.
- CPAF is more common in services contracting, especially in logistics services.

---

[21]Contracts can employ different contract types for different elements of the contract. The dominant contract type is recorded in FPDS-NG and reflected in these plots.

- CPIF and FPIF contracts with predetermined formulas that penalize for cost growth are most common in vehicle and weapon production as well as some in R&D, equipment related, and medical services.

The predominant contract types have shifted over the decades as policies intended to improve acquisition performance have emphasized certain types in certain situations. Figure 3-9 illustrates these changes by plotting the share of contract count in an historical dataset of 440 major MDAP development contracts from 1970 to 2010. Of note, CPAF contracting was emphasized in the 1990s but is being replaced by an emphasis on predetermined, formulaic cost incentives in CPIF (and, to a lesser extent, FPIF) contracts. Interestingly, CPFF dropped from its historical dominance in the 1970s to about 10–15 percent in this sample. This level is much lower than the larger dominance seen in Figure 3-8 across all DoD R&D service contracts.

Figure 3-10 shows the shifts in contract types for early production contracts in our dataset of 443 major MDAP contracts before 2010. Here we see that FPIFs were the dominant contract type for early production except in the 1990s, when hybrids and CPIF contracts surged. Trends from 2005 to 2010 showed a resurgence of FPIF and continuing use of CPIF but also an upward tick for CPAF. The CPAF data point is not reflected in the current FY13 obligation data, which reflects the current guidance that CPAF are intended only for cases when the government cannot predefine set performance criteria.

Figure 3-7. Contract Types Used Across All DoD Contracts (by dollars obligated; FY2013)
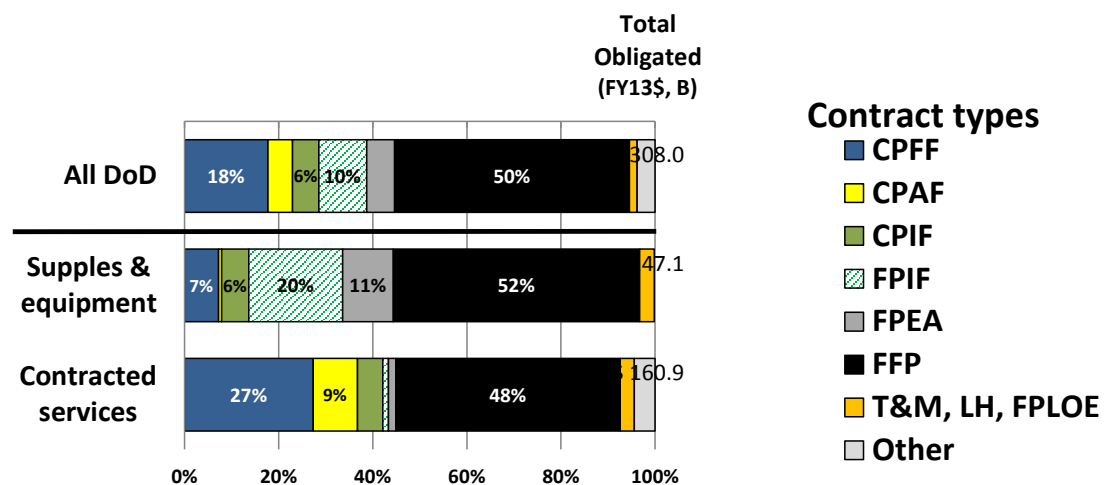
NOTE: Obligations are in billions of FY13 dollars.

Figure 3-8. Contract Types Used Across Portfolios (by dollars obligated; FY2013)

Supplies and equipment

	Total Obligated (FY13$, B)
Aircraft, Naval, and Land Vehicles	$ 53.9
Weapons & Ammunition	$ 18.0
Electronic & Communication Equipment	$ 19.4
Sustainment	$ 38.8
Facilities	$ 6.4
Clothing, Textiles & Subsistence	$ 10.5
Miscellaneous	$ 0.1

Aircraft, Naval, and Land Vehicles: 11%, 7%, 44%, 30%
Weapons & Ammunition: 7%, 19%, 11%, 55%
Electronic & Communication Equipment: 7%, 86%
Sustainment: 9%, 36%, 48%
Facilities: 86%
Clothing, Textiles & Subsistence: 95%
Miscellaneous: 100%

Contracted services

	Total Obligated (FY13$, B)
Research and Development	$ 28.1
Knowledge Based	$ 34.4
Logistics Management	$ 8.4
Equipment Related	$ 16.1
Electronic & Communication	$ 16.2
Medical	$ 13.3
Transportation	$ 8.7
Facility	$ 23.8
Construction	$ 12.0

Research and Development: 47%, 18%, 11%, 13%
Knowledge Based: 36%, 43%, 9%
Logistics Management: 26%, 47%, 24%
Equipment Related: 21%, 12%, 14%, 42%
Electronic & Communication: 27%, 58%
Medical: 56%, 14%, 7%, 22%
Transportation: 20%, 75%
Facility: 7%, 80%
Construction: 96%

NOTE: Obligations are in billions of FY13 dollars.

Figure 3-9. Contract Type Historical Trends on Major Development MDAP Contracts (1970–2010)
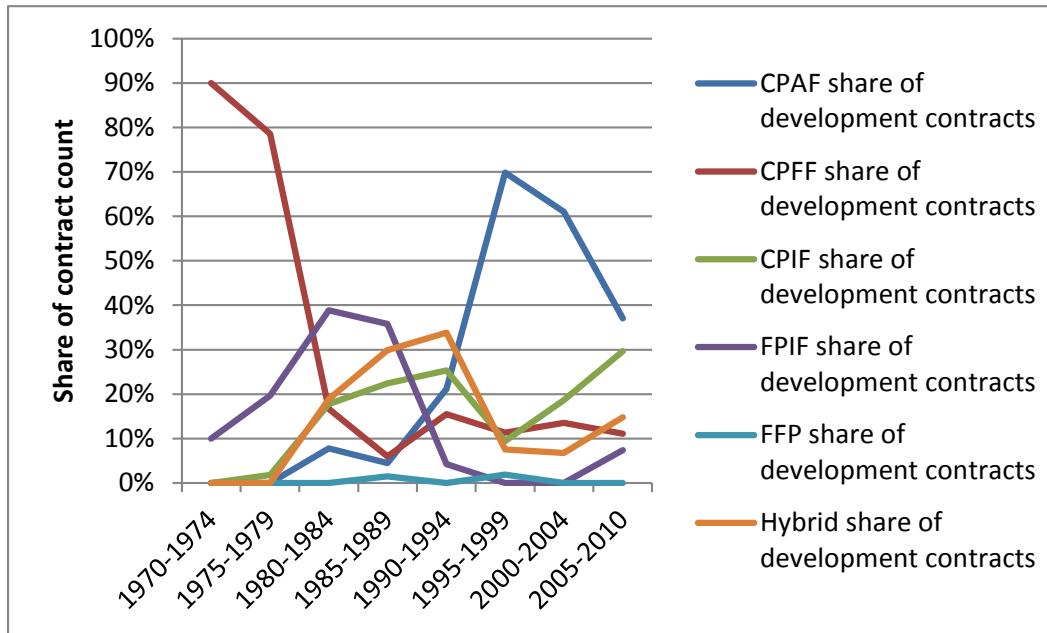
NOTE: *n*=440 development contracts. These contracts are primarily a subset of the "Research and Development" contracted services portfolio.

Figure 3-10. Contract Type Historical Trends on Major Early Production MDAP Contracts (1970–2010)
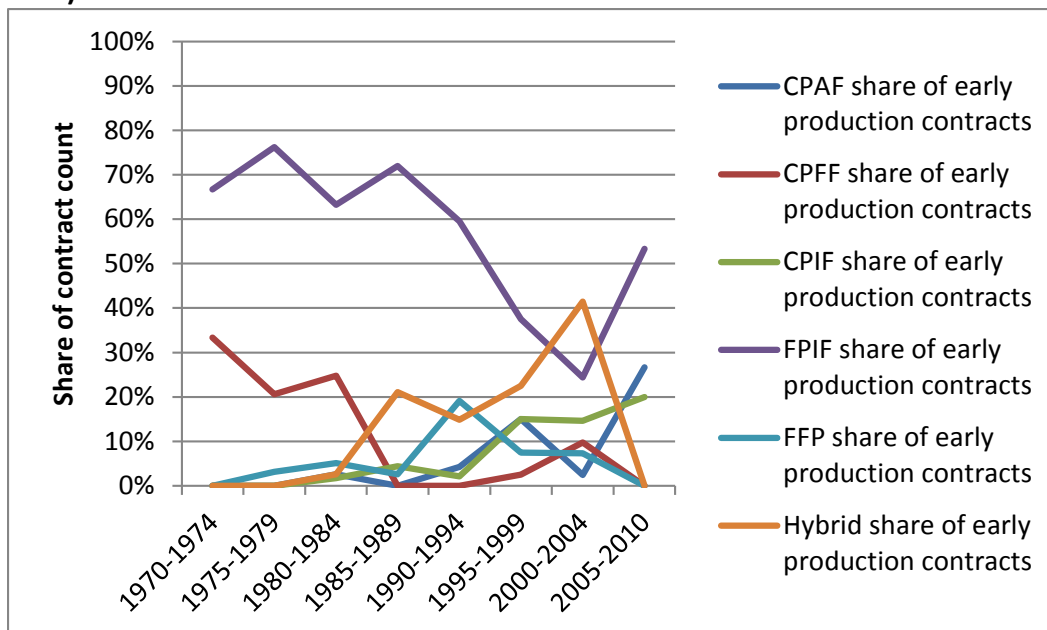
NOTE: *n*=433 early production contracts. These contracts are primarily a subset of the contracts in the "Aircraft, Ships/Submarines & Land Vehicles," "Weapons & Ammunition," and "Electronic & Communication Equipment" portfolios in supplies and equipment.

COMPETITION EFFECTS ON CONTRACTOR PERFORMANCE

As noted above, competition has long been a mainstay in both the public and commercial sectors for motivating contractors to provide the best value (i.e., the best performance at the lowest price). Competition has been deemed so valuable that we have set a strategic objective of increasing the percentage of spend on competed contracts from current levels (see Table 2-1 above).

We examined the degree to which prime MDAP contracts that were competed with at least two bidders showed better cost-, price-, and schedule-growth performance and how competition affected final margins. We examined sole-source and one-bidder contracts for new systems separately from modification upgrades to existing systems to determine whether modification (wherein the original manufacturer generally has a competitive edge in understanding the design) was a meaningful distinction.

In development, about two-thirds of the contracts in our MDAP contract dataset (both by count and by share of spend) were competed with at least two bidders. The remaining one-third was split about evenly between modification and new system contracts. As shown in Figure 3-11, price and schedule growth were lower on competed and sole-source modification contracts than on sole-source contracts for new systems. *These analyses demonstrate the general benefit of competition on cost, price, and schedule control.* Like price growth, cost growth (not shown) for competed and sole-source modification contracts was lower than that for sole-source contracts for new systems.

This data cannot tell us why cost and price growth are lower on competed contracts. We might have expected competitors to bid below their cost estimates with the expectation of making up any loses in engineering cost changes, but that does not appear to be happening systematically. It might be that competition is driving bidders to conduct more careful cost analysis and thus more realistic bids. It might also be that contractors put their best people on competitive efforts and thus have better performance. Further analysis is needed.

Final margins on competed contracts were lower and had more variation than either type of sole-source contract. This result tends to support the postulate that contractors may be willing to forgo some near-term profits if the contractors determine the longer-term benefits from "winning the franchise" form a good investment. Both types of sole-source contracts had very little variation in final margin (at least half of the population was very close to the median of about 8 percent margin).

In production, only two of the 83 contracts in our dataset involved competition (one with only one bidder, and one with two or more bidders and, interestingly, an award protest). Thus, we had insufficient data to examine the effect of competition on performance in production.

This analysis does not measure the effect of competition on the absolute value of initial (or final) price but rather on price (and cost) growth relative to initial values.

Figure 3-11. Effects of Competition on Contract Performance and Final Margin for New and Upgraded MDAPs (2000–2013)
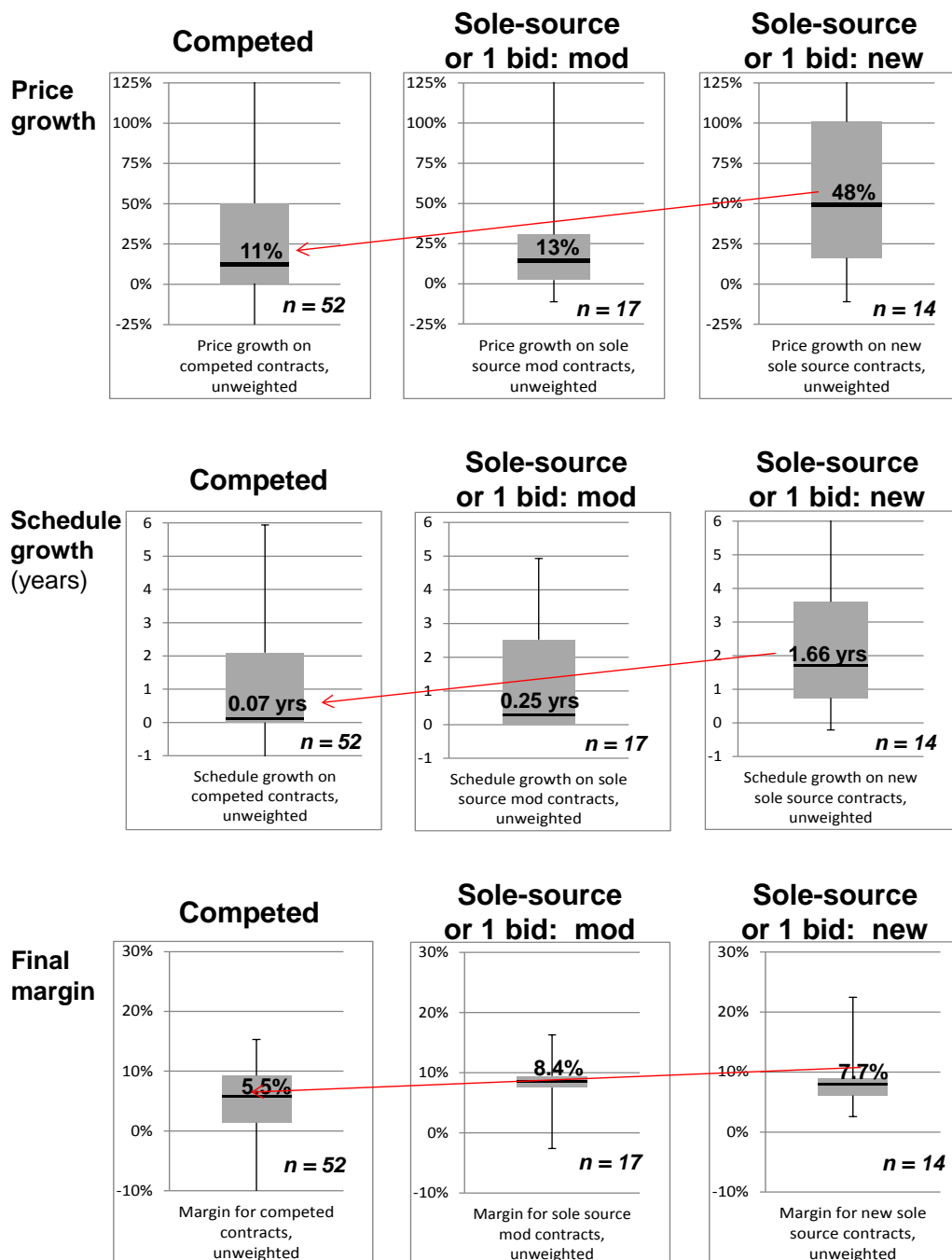
NOTE: mod = modification. Cost-growth distributions were similar to those for price growth. Price and cost growth were adjusted for inflation. Results are on a contract basis (i.e., not weighted by the dollar size of each contract). Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

PROGRAM MANAGER TENURE AND COST GROWTH

The acquisition workforce is a key contributor to the performance of the defense acquisition system. Unfortunately, existing data are often not readily available to examine which workforce factors correlate with acquisition performance. We are currently pursuing data that would allow such analyses (e.g., data that tie individuals to the programs they work on).

While larger data collections are being pursued, we are examining selected topics using manually collected data and existing program performance. So far we have examined whether program manager (PM) tenure correlates with program performance. Qualitatively, it has been asserted that having a stable person as PM should lead to better program performance.

There has been much discussion and some policymaking on the length of time that individuals are PMs. Despite vocal concerns about PM tenures being too short and policies that set a minimum tenure, analysis to date has not shown a correlation between PM tenure and program performance.

Relationship Between PM Tenure and Program Performance

In one relatively recent study, Ferreiro (2012) reviewed quantitative and qualitative studies, concluding that the performance of programs, as measured by unit-cost growth or Nunn-McCurdy Breaches, does *not* correlate with PM tenure. We examined PM tenure and two MDAP cost growth measures (total RDT&E and quantity-adjusted procurement). Regressions showed that practically none of the variation in cost growth was explainable by PM tenure.

We do not yet know why PM tenure does not correlate more strongly with program outcomes. It may be that the length of a PM's tenure matters less than the skills of the PM or of the ongoing staff that supports the PM (e.g., deputy, chief engineer, and contracting officer). Also, keeping an ineffective PM in place longer would logically make matters worse. We simply do not know analytically without controlling for other variables. Alternatively, it could be that, within reason, PM tenure alone is not critical in the way we staff program management teams. We are pursuing new data to link existing workforce databases to programs so we can examine various correlations, such as how the entire program leadership team tenure, experience, and background relate to program outcomes. We are convinced that superior management quality does matter—and matters a great deal—but correlating that belief to variables we can control is more difficult.

(This page was left blank intentionally)

4. CONCLUSIONS AND NEXT STEPS

Generally, we continue to see some signs of downward cost, price, and schedule growth, but there is ample room for continued improvement. The Department's continuous improvement efforts (including Better Buying Power and other activities) are ongoing. Only through steady pursuit of basic principles, informed and improved through analysis such as that presented in this report, do we expect to continue to improve acquisition.

In terms of the technical performance of our acquired systems, the readily available DOT&E reports show relatively flat trends in recent years. Of course, the real test is how effective our warfighters are on the battlefield when using the capabilities we acquire. There the United States continues to dominate, although we have increasing concerns about how well we are competing against evolving threats to our technological superiority as we continue to cope with budget reductions and uncertainty.

CONCLUSIONS

Combined with the results from last year's report, these analyses have produced a number of insights we are acting upon.

Not all incentives work. Contractual incentives are effective if (1) we use them; (2) they are significant, stable, and predictable; and (3) they are tied directly to our objectives.

"Cost-plus versus fixed-price" is a red herring. The distinction between cost-plus and fixed-price contracts *is not the divide on effectiveness*. Rather, the emphasis should be on matching incentives to the situation at hand instead of expecting fixed-price contracting to be a magic bullet. Fixed-price contracts have lower costs *because* they are used in lower-risk situations, not because they control costs better. Moreover, prices on fixed-price contracts are only "fixed" if the contractual work content and deliverables remain fixed, which is often not the case. Our analysis showed that objectively determined incentives were the factors that controlled costs, not selecting cost-plus or fixed-price contract types.

CPIF and FPIF contracts perform well and share realized savings. These contract types control cost, price, and schedule as well as, or better than, other types—and with generally lower margins. We pay for the technical risks on our developmental systems—unlike the private sector, where companies pay for R&D on new products. This is partly due to the fact that we are, to some degree, the only customer for new military products (i.e., a monopsony-type market). Thus, it makes sense to use incentives that (1) link profit to performance, (2) control price, and (3) share in cost savings, especially in production when the risks are low. Specific

incentive structures may not be appropriate in certain cases, so professional judgment is needed as always in matching contract type and incentives to the desired outcome.

FFP contracting requires knowledge of actual costs. FFP contracts provide vendors a strong incentive to control costs, especially in production, where they are most common. However, taxpayers do not share in those cost savings, unless the negotiated price took into account actual prior costs and margins, as well as the contractor's anticipated ability to continue cost reduction. Thus, to use FFP contracts effectively, we must fully understand actual costs when negotiating subsequent production lots.

Competition is effective—when viable. Competed contracts perform better on cost, price, and schedule growth than new sole-sourced or one-bidder contracts in development. (i.e., a contractor's knowledge that a competitor is not available may affect bidding and subsequent performance relative to that bid.) Thus, we must continue our efforts to seek competitive environments in creative ways. Unfortunately, direct competition on some MDAP contracts is often not viable—especially in production, where significant entry costs, technical data rights, or infrastructure may be barriers. In response, we are seeking ways in which competitive environments and open-system architectures will allow us to introduce competitive pressures.

Production margins may help minimize development time. Our analysis indicates that the prospect of higher margins in production may motivate contractors to complete development as soon as possible. Unfortunately, this assertion is hard to test quantitatively given the predominance of higher margins in all production contracts. However, we did find significant examples where production margins were smaller when development schedules slipped despite an explicit policy to this effect. As specified under the Better Buying Power initiatives, we need to continue creating a direct link between margin and performance, including getting into production as soon as possible. As our analysis indicates, competitive pressures also help to control schedule growth in development.

IMPROVING ACQUISITION PERFORMANCE THROUGH ANALYSIS

Continuous improvement is the essence of the Better Buying Power concept. Some efforts are very active and emphasized in public while others are either ongoing or in preparation. Below are a few additional topics related to acquisition performance for which we have some insights and face some analytic challenges.

Should-Cost

We are continuing the effort to change the acquisition culture from one focused on obligation rates, spending available budget, and accepting costs as a given—to one where managers scrutinize each element of cost under their control and assess how it can be reduced without reductions in value received. Under our "Should-Cost" initiative, specific efforts are executed to drive out cost, and these vary depending on the nature of the program and stage of the life cycle. The Department applies the savings elsewhere in the same program (helping to

incentivize even further should-cost efforts) or elsewhere in the Component. Nearly all DoD acquisition programs at all levels now have should-cost initiatives under consideration. We expect to provide data reflecting the impact of should cost implementation in our next report.

Cycle Time

Research is continuing in an effort to understand the degree to which cycle time is increasing between programs and, if so, what causes these increases. Generally, the analysis from last year's report indicates that cycle time on development contracts has increased by about one-sixth (0.9 years on a base of 5.2 years) since 1980. However, the variation in cycle time among programs is much larger than this temporal trend, and the complexity and capability of our weapons systems have increased dramatically during this time. Thus, cycle time appears longer compared with that of many decades ago, but the real driver appears to be system complexity. Moreover, some outliers are enormous; these probably represent more of a problem than the general increasing trend, again due to complexity. Ongoing analysis aims to improve our understanding of cycle time, when it is a problem, and how to address it in these cases.

Performance on Contracted Services

Fifty-three percent of all defense contracts in FY13 were for services rather than goods (see Figure 1-5 in Chapter 1). While much of the focus of defense acquisition has been on program acquisition (especially MDAPs), it is equally important to know how well acquired services are performing and what steps we can take to improve performance. Thus, as part of the Better Buying Power continuous improvement efforts we have been expanding the policies and structures related to the acquisition of services. Part of those efforts involves identifying and institutionalizing performance measures that management can use to understand service performance. This effort will take time, but we continue to look for objective data to inform the performance of these contracts. Currently, this includes the metrics on competition and small-business rates reported earlier.

Workforce and Performance

Another area important to defense acquisition relates to the acquisition workforce. Apart from the qualitative Procurement Management Reviews (overseen by the Defense Contract Management Agency) and the Procurement Management Reviews/Assessment (conducted by the military departments), we are seeking data to link our data on individuals in our human capital databases to the programs and activities that they perform. When available, this will allow us to look for statistical correlations between workforce factors (e.g., technical background, certification levels, and experience) and outcomes on programs and activities. The earlier investigation of PM tenure summarized in this report provides an example of this type of analysis, but we need broader data to help control for multiple variables rather than just examining a single variable such as tenure against program performance.

THE FUTURE

Acquisition performance is about continuous improvement across all phases, functions, and professional specialties. There is no end state. We will continue to build on the analytical work completed so far to go on separating fact from myth about the most effective way to improve defense acquisition outcomes.

A. STATISTICAL DETAILS

The statistical analyses conducted involved both parametric and nonparametric tests.

Supporting Sample Analysis for Regressions

In our linear multivariate regression analyses, we conducted supporting sample analysis tests for normality of residuals (Smirnov-Kolmogorov and Shapiro-Wilk tests), heteroskedasticity (Cook-Weisberg test), multicollinearity (variance inflation factor test), omitted variables (Ramsey Regression Equation Specification Error Test [RESET]), and correct model specification (Linktest). We also used bootstrap simulations to obtain unbiased coefficient estimates, correct standard errors, and correct confidence intervals.

Single Variable Analysis Tests

Single variable analyses allowed us to focus on differences by a single factor (e.g., phase, contract type, cost or price growth, schedule growth, or final margin). Nonparametric tests (Wilcoxon rank-sum and Kolmogorov-Smirnov) were used to test for statistical significance between populations, and the median was used as the measure of central tendency because the distributions were skewed.

Eliminating Outliers in Trend Analysis

Extreme outliers can overly bias trend analysis, leading to trends that are too dependent on one or two significant outliers. As a result, the following four tests were employed on our analysis of cost-growth, price-growth, schedule-growth, and final-margin trends:
- Studentized residual test
- Leverage test
- Cook's Distance Test
- DFITS (Difference in Fits, Standardized) Test [also called DFFITS]

For example, our MDAP dataset that includes price and margin had a small sample of major programs from before year 2000. These tests determined that those temporal outliers (along with a small number of others) should be rejected in our trend analysis.

Comments on Interpreting Box-and-Whisker Diagrams

Throughout this report, the so-called "box and whisker" charts (see Figure A-1) help visualize the distribution of a particular variable. The grey boxes show the second and third quartiles (i.e., the 25th to the 50th percentile, and the 50th to the 75th percentile). The minimum and maximum are shown with a small bar at the end of the vertical line (or may run off the chart in

some instances). The median (50th percentile, where half of the occurrences are above it and half below) is the *best measure of central tendency* in the data because the distributions are skewed. Note that the quartiles do not convey the actual distributions within the quartiles. As seen by the illustrations on the left of the figure, these distributions can be "lumpy" or nonuniform, but the charts do provide a quick visual for comparing two distributions. The charts also convey a sense of how much of the distribution is, say, negative or larger than a value of interest.

Figure A-1. Key to Reading the "Box and Whisker" Charts

(illustration)
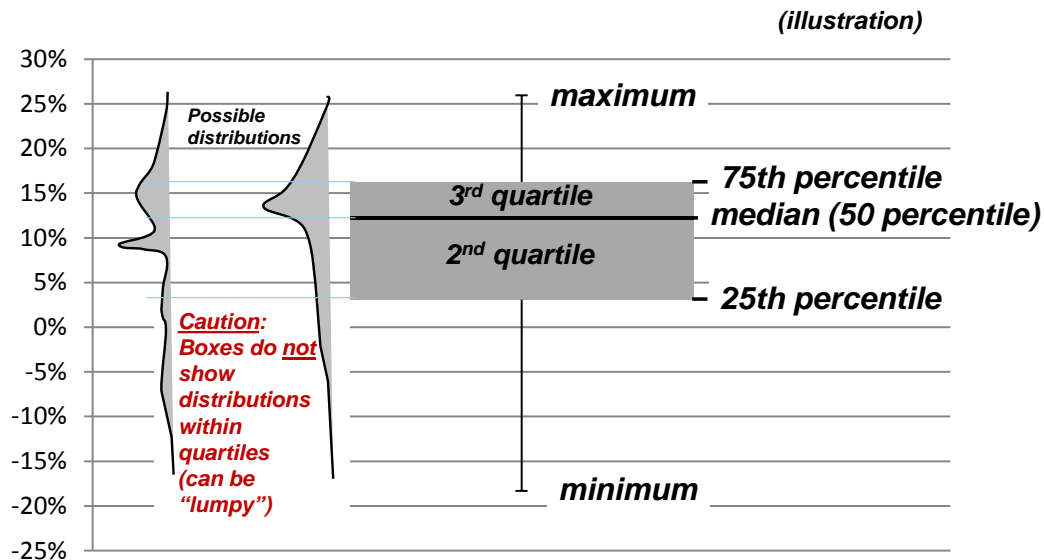
B. ABBREVIATIONS

AAC—Air Armament Center [Air Force]

ABC—AmerisourceBergen

AEHF—Advanced Extremely High Frequency

AFATDS—Advanced Field Artillery Tactical Data System

AFMC—Air Force Materiel Command

AMC—Army Materiel Command

AMCOM—Aviation and Missile Command [Army]

AMP—Avionics Modernization Program

AoA—Analysis of Alternatives

APUC—Average Procurement Unit Cost

ARH—Armed Reconnaissance Helicopter

ASC—Aeronautical Systems Center [Air Force]

AT&L—Acquisition, Technology, and Logistics

BA—Boeing

BAMS—Broad Area Maritime Surveillance

BY—base year

CAE—Component Acquisition Executive

CAH—Cardinal Health

CBB—Contract Budget Base

CEDD—Continuing Engineering Design/ Development

Chem Demil—chemical [weapons] demilitarization

CJR—Cobra Judy Replacement

CP—cost plus

CPAF—cost-plus-award-fee

CPFF—cost-plus-fixed-fee

CP/H—cost-plus and hybrid

CPIF—cost-plus-incentive-fee

CY—constant year; calendar year

DCMO—Office of the Deputy Chief Management Officer

DDS—Dry Dock Shelter

DFITS—Difference in Fits, Standardized [also called DFFITS]

DoD—Department of Defense

DOT&E—Director, Operational Test and Evaluation

DPAP—Defense Procurement and Acquisition Policy

DSUP—Defense System Upgrade Program

DVH—Double-V Hull

EBIT—earnings before interest and taxes

EBITDA—earnigs before interest, taxes, depreciation, and amortization

EELV—Evolved Expendable Launch Vehicle

EFV—Expeditionary Fighting Vehicle

ELX—Excelis

EMD—Engineering, Manufacturing and Development

ESRX—Express Scripts

FAB-T—Family of Advanced Beyond-line-of-sight Terminals

FAR—Federal Acquisition Regulation

FCS—Future Combat Systems

FDD—Full-Deployment Decision

FOC—Full operational capability

FOIA—Freedom of Information Act

FP—fixed price

FPDS-NG—Federal Procurement Data System-Next Generation

FPIF—fixed-price incentive firm [target]

FFO—[date] funds first obligated

FFP—firm-fixed-price

FFRDC—Federally Funded Research and Development Center

FP—fixed price

FPDS—Federal Procurement Data System

FOC—full operational capability

FY—Fiscal Year

GD—General Dynamics

Granite Sentry—aircraft tracking system

GBR—Ground-Based Radar

GBS—Global Broadcast Service

GCCS—Global Command and Control System

GCSS—Global Combat Support System

GPS—Global Positioning System

HII—Huntington Ingalls Industries

HP—Hewlett Packard

IAMD—Integrated Air and Missile Defense

IR&D—Independent Research and Development

JASSM—Joint Air-to-Surface Stand-off Missile

JDAM—Joint Direct Attack Munitions

JHSV—Joint High Speed Vessel

JLENS—Joint Land Attack Cruise Missile Defense Elevated Netted Sensor System

JM&LLCMC—Joint Munitions and Lethality Life Cycle Management Command

JSF—Joint Strike Fighter

JSOW—Joint Stand-off Weapon

JSTARS—Joint Surveillance Target Attack Radar System

JTRS—Joint Tactical Radio System

KPP—Key Performance Parameter

LADS—Laser Area Defense System

LMC—Lockheed Martin

LRIP—Low-Rate Initial Production

MAIS—Major Automated Information Systems

MANT—ManTech

108

MAR—MAIS Annual Report

MCK—McKesson

MDA—Milestone Decision Authority

MDAP—Major Defense Acquisition Program

MDD—Materiel Development Decision

MEADS—Medium Extended Air Defense System

MIDS-LVT—Multi-functional Information Distribution System—Low Volume Terminal

MilCon—Military Construction

MilPer—Military Personnel

MRAP—Mine Resistant Ambush Protected

MS—Milestone

MUOS—Mobile User Objective System

NAICS—North American Industry Classification System

NAVAIR—Naval Air Systems Command

NAVSEA—Naval Sea Systems Command

NDAA—National Defense Authorization Act

NGC—Northrop Grumman

O—other

O&M—Operations and Maintenance

OCO—Overseas Contingency Operations

PAC-3—Patriot Advanced Capability–3

PAD—Preferred Alternative Decision

PARCA—Performance Assessments and Root Cause Analysis

PAUC—Program Acquisition Unit Cost

PM—program manager

PM EAC—Program Manager Estimate at Completion

PSC—product service code

R&D—research and development

RDT&E—Research Development Test and Evaluation

RESET—Ramsey Regression Equation Specification Error Test

RGTS—Al Raha Group for Technical Services

RTN—Raytheon

S&E—supplies and equipment

SAE—Service Acquisition Executive

SAIC—Science Applications International Corp.

SAR—Selected Acquisition Report

S&T—science and technology

SBIRS—Space-based Infrared System

SCAMP—Single-channel Anti-jam Man-portable

SEC—[U.S.] Securities and Exchange Commission

TMD—Theater Missile Defense

TY—then year

UCA—Undefinitized Contract Action

UT—United Technologies

ULA—United Launch Alliance

U.S.—United States

U.S.C.—U. S. Code

USD—Under Secretary of Defense

USD(AT&L)—Office of the Under Secretary of Defense, Acquisition, Technology, and Logistics

WIN-T—Warfighter Information Network–Tactical

WLDAC—World Airways.

YPUC—Yearly Production Unit Cost

C. Indices

Index of Figures

INDEX OF TABLES

(This page was left blank intentionally)

D. REFERENCES

Arena, Mark V., Irv Blickstein, Abby Doll, Jeffrey A. Drezner, James G. Kallimani, Jennifer Kavanagh, Daniel F. McCaffrey, Megan McKernan, Charles Nemfakos, Rena Rudavsky, Jerry M. Sollinger, Daniel Tremblay, and Carolyn Wong, *Management Perspectives Pertaining to Root Cause Analyses of Nunn-McCurdy Breaches, Volume 4: Program Manager Tenure, Oversight of Acquisition Category II Programs, and Framing Assumptions*, Santa Monica, Calif.: RAND Corp., MG-1171/4-OSD, 2013. http://www.rand.org/pubs/monographs/MG1171z4.html

Arena, Mark V., Robert S. Leonard, Sheila E. Murray, and Obaid Younossi, *Historical Cost Growth of Completed Weapon System Programs*, Santa Monica, CA: RAND Corp., TR-343-AF, 2006. http://www.rand.org/pubs/technical_reports/TR343.html

Deputy Chief Management Officer (DCMO), *Strategic Management Plan: The Business of Defense FY2014–FY2015*, U.S. Department of Defense, July 1, 2013.

Ferreiro, Larrie D., "How Does Program Manager Tenure Affect Program Performance? A review of the findings to date,"unpublished, 2012.

Flyvbjerg, Bent, Mette Skamris Holm, and Søren Buhl, "Underestimating Costs in Public Works Projects: Error or Lie?" *Journal of the American Planning Association, 68*(3): 279–295, 2002.

Hough, Paul G., *Pitfalls in Calculating Cost Growth from Selected Acquisition Reports*, Santa Monica, Calif.: RAND Corp., N-3136-AF, 1992. http://www.rand.org/pubs/notes/N3136.html.

Jarvaise, Jeanne M., Jeffrey A. Drezner, and Daniel M. Norton, *The Defense System Cost Performance Database: Cost Growth Analysis Using Selected Acquisition Reports*, Santa Monica, Calif.: RAND Corp., MR-625-OSD, 1996. http://www.rand.org/pubs/monograph_reports/MR625/.

Laffont, Jean-Jacques, and David Martimort, *The Theory of Incentives: The Principal-Agent Model*, Princeton, NJ: Princeton Univ. Press, 2001.

McNicol, D. L., *Growth in the Costs of Major Weapon Procurement Programs, 2nd Edition*, Alexandria, VA: Institute for Defense Analyses, IDA Paper P-3832, 2005.

Schwartz, Moshe, *The Nunn-McCurdy Act: Background, Analysis, and Issues for Congress*, Congressional Research Service, June 21, 2010.

Thompson, Thomas, Jr., Ken Cameron, and Addie Adkins, *Financial Executive Compensation Survey 2013*, Morristown, NJ: Financial Executives Research Foundation, 2013.

Under Secretary of Defense (Acquisition, Technology, and Logistics), *Performance of the Defense Acquisition System, 2013 Annual Report,* U.S. Department of Defense, Washington, DC, June 28, 2013.

U.S. General Services Administration, *Product and Service Codes Manual—Federal Procurement Data System*, Office of Governmentwide Policy, Aug. 2011.

Younossi, Obaid, Mark V. Arena, Robert S. Leonard, Charles Robert Roll, Jr., Arvind Jain, and Jerry M. Sollinger, *Is Weapon System Cost Growth Increasing? A Quantitative Assessment of Completed and Ongoing Programs*, Santa Monica, Calif.: RAND Corp., MG-588-AF, 2007. http://www.rand.org/pubs/monographs/MG588.html